

Access DB#

54659

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER-DISK-E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers; and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

**STAFF USE ONLY****Type of Search****Vendors and cost where applicable**

Searcher: \_\_\_\_\_ NA Sequence (#) 2 STN \_\_\_\_\_  
 Searcher Phone #: 4498 AA Sequence (#) 6 Dialog \_\_\_\_\_  
 Searcher Location: \_\_\_\_\_ Structure (#) \_\_\_\_\_ Questel/Orbit \_\_\_\_\_  
 Date Searcher Picked Up: 11/9 Bibliographic \_\_\_\_\_ Dr. Link \_\_\_\_\_  
 Date Completed: 11/13 Litigation \_\_\_\_\_ Lexis/Nexis \_\_\_\_\_  
 Searcher Prep & Review Time: \_\_\_\_\_ Fulltext \_\_\_\_\_ Sequence Systems ✓  
 Clerical Prep Time: 20 Patent Family \_\_\_\_\_ WWW/Internet \_\_\_\_\_  
 Online Time: 15 Other \_\_\_\_\_ Other (specify) \_\_\_\_\_



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:30:16 ; Search time 79.75 Seconds

(without alignments)  
161.917 Million cell updates/sec

Title: US-09-668-021-2

Sequence: 1 MQLPLALCVLCVHTAFRV.....KKPRARSAKANAQAELENAV 213

Scoring table:

Gapop 60.0 , Gapext 60.0

Archived: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

A\_Geneseq.0601.\*  
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	100.0	213	21	Human DAN/Cerberus
2	213	100.0	213	21	Human TGF-beta bin
3	213	100.0	213	22	Human secreted pro
4	175	82.2	213	21	Human TGF-beta bin
5	175	82.2	213	21	Human TGF-beta bin
6	139	65.3	213	21	Human DAN/Cerberus
7	135	63.4	213	21	Human TGF-beta bin
8	76	33.7	176	21	Human TGF-beta bin
9	49	23.0	50	20	Human TGF-beta bin
10	33	15.5	213	21	Human 5' EST secre
11	32	15.0	211	21	Murine TGF-beta bi

12	23	10.8	23	21	AAV96435
13	9	4.2	116	18	AAW27654
14	9	4.2	116	18	AAW44090
15	9	4.2	206	18	AAW09408
16	9	4.2	206	18	AAW58704
17	9	4.2	206	20	AAW95711
18	9	4.2	206	21	AAW10233
19	9	4.2	206	21	AAW75981
20	9	4.2	206	21	AAW76031
21	9	4.2	206	22	AAW55920
22	9	4.2	206	22	AAW55970
23	8	3.8	428	20	AAV52704
24	8	3.8	428	20	AAV32824
25	7	3.3	61	21	AAW00348
26	7	3.3	114	19	AAW68587
27	7	3.3	114	19	AAW75969
28	7	3.3	173	21	AAW43980
29	7	3.3	177	20	AAV33001
30	7	3.3	186	19	AAW65018
31	7	3.3	219	21	AAW42285
32	7	3.3	265	21	AAW42284
33	7	3.3	282	21	AAW42283
34	7	3.3	285	21	AAW29625
35	7	3.3	412	16	AAW74207
36	7	3.3	412	19	AAW71369
37	7	3.3	412	20	AAW06478
38	7	3.3	412	21	AAV93685
39	7	3.3	426	22	AAW64892
40	7	3.3	554	20	AAV36226
41	7	3.3	718	19	AAW72072
42	7	3.3	917	18	AAW37437
43	7	3.3	917	18	AAW37429
44	7	3.3	917	18	AAW23793
45	7	3.3	922	20	AAW33002
46	7	3.3	1064	22	AAW46588
47	7	3.3	1064	22	AAW46589
48	6	2.8	10	12	AAW10268
49	6	2.8	15	19	AAW45818
50	6	2.8	15	19	AAW45813

#### ALIGNMENTS

RESULT 1	
ID	AAW26106 standard; Protein; 213 AA.
XX	
AC	AAW26106;
XX	
DT	15-JAN-2001 (first entry)
XX	
DE	Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
XX	
KW	Human; DAN/Cerberus-related protein 6; hDCR6; morphogenic protein; antagonist; BMP; cell growth; cell differentiation; bone formation; gene therapy.
KW	
OS	Homo sapiens.
XX	
PN	W0200055193-A2.
XX	
PD	21-SEP-2000.
XX	
PF	02-MAR-2000; 2000MO-US05537.
XX	
PR	12-MAR-1999; 99US-024118.
XX	
PA	(REGE-) REGENERON PHARM INC.
XX	
PI	Economides AN;
XX	
DR	WPI; 2000-638179/61.

Mutant human TGF-b  
Secreted protein A  
Human secreted pro  
Human small CCN-1i  
Human small CCN-1i  
Homo sapiens fetal  
Human adult retina  
Murine skin cell p  
Murine skin cell p  
Skin cell protein,  
Human ELK-1 protei  
Human ELK-1 protei  
Human secreted pro  
Tobacco hawkmoth E  
Ecodysis tritigerling  
Human cancer assoc  
Human serine prote  
Forsythia dirigent  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Cat flea HMT perit  
Human death associ  
Death associated p  
Human tumour-assoc  
Amino acid sequenc  
Human secreted pro  
Human secreted pro  
H5V-2 strain SB5 C  
Rat hexokinase II.  
Rat hexokinase II.  
AS-30D tumour type  
Human serine prote  
Serine protease in  
Serine protease in  
Human ventricular  
Peptide recognised

DR N-PSDB: AAA94051.  
XX Novel isolated, human DNA/Cerberus related protein 6 which include  
PT natural homologue, and polypeptides comprising DCR6 domain and nucleic  
PT acids encoding the proteins which are useful as probes and primers -  
XX  
PS Claim 8; Fig 3; 40pp; English.  
XX  
CC The present sequence comprises the amino acid sequence encoded by exons 1  
CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding  
CC sequence. The coding sequence was isolated from a human kidney cDNA  
CC library containing exons 1 and 4 of the sequence. HPCR6 is closely  
CC related to the DAN and DCR5 proteins, both of which act as antagonists of  
CC morphogenic proteins such as BMP. It is possible that the hPCR6 gene and  
CC protein can be used as immunogens, modulators of cell function, growth  
CC and differentiation, to reduce undesirable bone formation, to identify  
CC DCR6 binding agents, in diagnosis, and in gene therapy.  
XX  
SQ Sequence 213 AA;

Query Match 100.0%; Score 213; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 2,1e-212;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOLPLALCLVCLVHTAFRVVEGOGMOAFKNDATETIPELGSEYPEPPPELNNKTMNRAE 60  
DB 1 mqlplalclvclvhtafrrveggwqafkndatetlpelegseypepppelelnkkmtrae 60  
OY 61 NGGRPPHHPEFTKDVSEYSCRELHFTRYVDGFCRSARKPYTELVCSGCGPARLLPNAIG 120  
DB 61 nggrpphhpeftkdvseycrelhfttryvdgpcrsarkpytelvcsgcgparllpnaig 120  
OY 121 RGMWRPSPGDFRCIPDRYRAQRYVLLCPGGEAPRARKVRLVASCKCKRLTRFHNSQLK 180  
DB 121 rgkmwrpsgpdfrcipdryraqrvllcpggeaprarkevrlvasckckrltrfhnselk 180  
OY 181 DFGTEAARPOKGRKPRPRARSAKANQAELENAY 213  
DB 181 dfgteaarppkgrkprprarsakangaelenay 213

RESULT 2  
AA96429  
ID AA96429 standard; Protein: 213 AA.  
XX  
AC AA96429;  
XX  
DT 12-SEP-2000 (first entry)  
XX

Human TGF-beta binding protein (BEER).  
XX  
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
XX

XX Homo sapiens.  
XX  
PM WO200032773-A1.  
XX  
PD 08-JUN-2000.  
XX  
PF 24-NOV-1999; 99MO-US27990.  
XX  
PR 27-NOV-1998; 98US-0110283.  
XX  
PA (DARW-) DARWIN DISCOVERY LTD.  
XX  
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
PI Van Ness J, Winkler DG;  
XX WPI: 2000-412321/35.  
DR N-PSDB: AAA29055.  
XX

PT Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
PS Claim 2; Page 116; 162pp; English.  
XX

CC This shows the human transforming growth factor-beta (TGF-beta)  
CC binding protein designated hBEER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patients own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other  
CC disorders associated with low mineral content.  
XX

Query Match 100.0%; Score 213; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 2,1e-212;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOLPLALCLVCLVHTAFRVVEGOGMOAFKNDATETIPELGSEYPEPPPELNNKTMNRAE 60  
DB 1 mqlplalclvclvhtafrrveggwqafkndatetlpelegseypepppelelnkkmtrae 60  
OY 61 NGGRPPHHPEFTKDVSEYSCRELHFTRYVDGFCRSARKPYTELVCSGCGPARLLPNAIG 120  
DB 61 nggrpphhpeftkdvseycrelhfttryvdgpcrsarkpytelvcsgcgparllpnaig 120  
OY 121 RGMWRPSPGDFRCIPDRYRAQRYVLLCPGGEAPRARKVRLVASCKCKRLTRFHNSQLK 180  
DB 121 rgkmwrpsgpdfrcipdryraqrvllcpggeaprarkevrlvasckckrltrfhnselk 180  
OY 181 DFGTEAARPOKGRKPRPRARSAKANQAELENAY 213  
DB 181 dfgteaarppkgrkprprarsakangaelenay 213

RESULT 3  
AA97589  
ID AA97589 standard; Protein: 213 AA.  
XX  
AC AA97589;  
XX  
DT 05-APR-2001 (first entry)  
XX  
DE Human secreted protein PRO7476.  
XX

XX Secreted protein: human; PRO protein; neoplastic cell growth; tumour;  
KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;  
KW angiogenic disorder; immunologic disorder; PRO7476.  
XX

XX Homo sapiens.  
XX  
PM WO200075317-A2.  
XX  
PD 14-DEC-2000.  
XX  
PF 15-MAY-2000; 2000MO-US13358.  
XX  
PR 09-JUN-1999; 99US-0138385.  
PR 20-JUL-1999; 99US-0144790.  
PR 03-AUG-1999; 99US-0146843.  
PR 10-AUG-1999; 99US-0148188.  
PR 17-AUG-1999; 99US-0149320.  
XX

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 21:20:08 ; Search time 6020.71 Seconds  
(without alignments)  
5911.477 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301

Sequence: 1 agagcctgctactggaag.....caatgatcatgaccgaaag 2301

Scoring table:

OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

GenEmbl:\*

1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_da3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_om:\*  
9: gb\_ov:\*  
10: gb\_pat1:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: gb\_da1:\*  
17: gb\_da2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_hum:\*  
21: em\_htgo\_inv:\*  
22: em\_htgo\_inv:\*  
23: em\_htg\_hum1:\*  
24: em\_htg\_hum2:\*  
25: em\_htg\_hum3:\*  
26: em\_htg\_hum4:\*  
27: em\_htg\_hum5:\*  
28: em\_htg\_hum6:\*  
29: em\_htg\_hum7:\*  
30: em\_htg\_hum8:\*  
31: em\_htg\_inv1:\*  
32: em\_htg\_inv2:\*  
33: em\_htg\_inv2:\*  
34: em\_htg\_inv2:\*  
35: em\_hum1:\*  
36: em\_hum2:\*  
37: em\_hum3:\*  
38: em\_hum4:\*  
39: em\_hum5:\*  
40: em\_hum6:\*  
41: em\_hum7:\*  
42: em\_in:\*  
43: em\_om:\*  
44: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_ro1:\*  
95: gb\_ro2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_da3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2301	100.0	2323	89	AF326739	AF326739 Homo sapi
2	2230	96.9	2329	9	AX056687	AX056687 Sequence
3	2163	94.0	2296	89	AF331844	AF331844 Homo sapi
4	2036	88.5	21501	89	AF326736	AF326736 Homo sapi
5	2036	88.5	94752	85	AC003098	AC003098 Homo sapi
6	1915	83.2	151780	72	AC055813	AC055813 Homo sapi
7	661	28.7	177744	75	AC073954	AC073954 Homo sapi
8	231	10.0	177744	75	AC073954	AC073954 Homo sapi

9	110	4 8	642	89	AF326742	AF326742 Cercopit
10	83	3 6	532	7	AF326738	AF326738 Bos tauru
11	69	3 0	638	94	AF326740	AF326740 Mus muscu
12	69	3 0	17423	94	AF326737	AF326737 Mus muscu
13	69	3 0	110000	73	AC068782.2	Continuation (3 of
14	69	3 0	200727	62	AC012296	AC012296 Mus muscu
15	69	3 0	211533	73	AC068807	AC068807 Mus muscu
16	68	3 0	674	94	AF326741	AF326741 Rattus no
17	17	1 7	40	9	AX056701	AX056701 Sequence
18	34	1 5	83844	60	AC008708	AC008708 Homo sapi
19	34	1 5	135961	66	AC021288	AC021288 Homo sapi
20	33	1 4	835	54	G39379	G39379 221626 Zeb
21	33	1 4	189430	77	AC090123	AC090123 Mus muscu
22	32	1 4	139072	62	AC011261	AC011261 Homo sapi
23	32	1 4	181143	79	AL158140	AL158140 Homo sapi
24	31	1 3	126136	82	AL590642	AL590642 Homo sapi
25	31	1 3	140606	79	AL353761	AL353761 Homo sapi
26	31	1 3	160628	73	AC067756	AC067756 Homo sapi
27	31	1 3	200516	62	AC011308	AC011308 Homo sapi
28	30	1 3	108893	93	AC080073	AC080073 Homo sapi
29	29	1 3	158457	76	HS07130E4	AL078582 Human DNA
30	29	1 3	159974	90	AL139008	AL139008 Human DNA
31	29	1 3	179146	81	AL450401	AL450401 Homo sapi
32	29	1 3	187109	75	AC074100	AC074100 Homo sapi
33	29	1 3	195048	80	AL356489	AL356489 Homo sapi
34	29	1 3	200582	80	AL356093	AL356093 Homo sapi
35	29	1 3	216293	97	HRCRCH89	AF029308 Homo sapi
36	29	1 3	296820	94	AF312033	AF312033 Mus muscu
37	28	1 2	199	8	AF114775	AF114775 Ictalurus
38	28	1 2	254	94	HS191VE1	Z16819 H. sapiens
39	28	1 2	395	54	G39626	G39626 Z6994 Zebra
40	28	1 2	407	7	CBE243714	AJ243714 Cryptopro
41	28	1 2	474	8	AF314679	AF314679 Delistes
42	28	1 2	714	5	AF189701	AF189701 Biomphala
43	28	1 2	721	54	G39227	G39227 220157 Zeb
44	28	1 2	3857	94	AF199380	AF199380 Mus muscu
45	28	1 2	24035	94	MMU251155	AJ251155 Mus muscu
46	28	1 2	42859	90	AL390210	AL390210 Human DNA
47	28	1 2	85637	90	AL391002	AL391002 Human DNA
48	28	1 2	90323	85	AC000105	AC000105 Homo sapi
49	28	1 2	100000	91	AP000081	AP000081 Homo sapi
50	28	1 2	100000	91	AP000081	AP000081 Homo sapi

ALIGNMENTS

RESULT	1	AF326739	2323 bp	mRNA	PRI	28-FEB-2001
LOCUS	AF326739	Homo sapiens	sclerostin	mRNA, complete cds.		
DEFINITION	AF326739	AF326739.1	GI:13161019			
VERSION	AF326739.1	GI:13161019				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REFERENCE						
AUTHORS						
TITLE						

JOURNAL	Submitted (07-DEC-2000)	Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA		
FEATURES		
source		
1..2323		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="17"		
/map="17q21"		
1..47		
48..689		
/codon_start=1		
/product="sclerostin"		
/protein_id="AK13454.1"		
/db_xref="GI:13161020"		
/translation="MOPLALICLLVHTAFVNBEGGQWAKNATETIPLEGYP		
EPPPELENKTNNAENGRRPHHPEFTKDVSEKSELHFTYVTDGPRKAPYTE		
LVCSGGCPAPLPLPNAIGRKMWMPSPDRCLPDYRRAQVLLCPGGAPPARVR		
LVASCKCKRLTRFHNOSELKDFGEARAPQGRKPRPRASAKANOAELENAY"		
690..2323		
BASE COUNT	590 a	635 c 614 g 484 t
ORIGIN		
3'UTR		
Query Match	100.0%	Score 2301; DB 89; Length 2323;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 2301; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
1	agagcctgtgtctactcgggaagtgtgcgtccctctctctgtgcgtgtgtacatgacgtccac	60
1	AGAGCCTGTGTCTACTCGGAAGTGTGGCTCCCTCTCTGTGTGTACATGACATGCCCTCAC	60
61	tggccctgtgtctgtcgtcgtcgtcgtgtgtacacacagcctcgtgtagtgaaggccag	120
61	TGGCCCTGTGTCTGTCGTCTGCTGCTGTGTACACACAGCCTTCGTGTGTGTGAGGCGCAGG	120
121	gtgtgcaagcgttcaagaaatgacacgaaatcattcccgagctcggagatgacccg	180
121	GTTGTGCAAGCGTTTCAAGAAATGATGACGCAAGAAATCATCTCCGAGCTCGAGAGTACCCCG	180
181	agcctccacccgagctgtgaaagaacaaagacatgaacccgggaggaagaaggagc	240
181	AGCCTCCACCCGAGCTGTGAAAGAACAAGACATGAAACCGGCGGAGAGGCGGCG	240
241	ctccaccacccctttagacccaagaacggtgtccgagtaacgtgtcgcgagctgact	300
241	CTCCACCACCCCTTTAGACCAAGACGTGTCCGAGTACGTGCGCGAGCTGCACCT	300
301	tcaccgctacgtgacgagtgtgcggtgtgcgagcgcgaacgcggtacacgagctgtgt	360
301	TCACCGCTACGTGACGAGTGTGCGGCTGCGCACGCGCAGCCGATGCGAGCTGTGT	360
361	gtctcggcagctgtgcggcggcgcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	420
361	GTCCTCGGCAGTGTGCGGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	420
421	ggcgaactagtgagcgcgacgtcctgcgtgcatcccgacacgtctacccgagcgcgtgc	480
421	GCGCACTAGTGTGAGCGCGACGTCTCGTGCAATCCCGACCGGTACCGCGCAGCGCTGC	480
481	agctgctgtgtcccggt	540
481	AGCTGCTGTGTCCCGT	540
541	gcaagtgaagcgtcaacccgcttcaacaacagctgtgagatcgaagacttgagagcg	600
541	GCAAGTGAAGCGCTCAACCCGCTTCAACAACAGTGTGAGACTTGAAGACTTGGGACCG	600
601	agcgcgtcgc	660
601	AGCGCGTCTGCG	660
661	accgagcgcgagctgtgagaaagcctactagagcccgccgcgcgcctcccaacgcgagc	720
661	ACCGAGCGCAGCTGTGAGAAAGCCTACTAGAGCCCGCGCGCGCGCGCGCGCGCGCGCG	720

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2001, 01:07:51 ; Search time 6020.71 Seconds  
(without alignments)  
5911.477 Million cell updates/sec

Title: US-09-668-021-5  
Perfect score: 2301  
Sequence: 1 agagcctgtgtactgtgaag.....caatgaatcactgaccgaag 2301

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

arched: 1344157 seqs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_rod:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_vil:\*  
59: gb\_vil2:\*  
60: gb\_vil3:\*  
61: gb\_vil4:\*  
62: gb\_vil5:\*  
63: gb\_vil6:\*  
64: gb\_vil7:\*  
65: gb\_vil8:\*  
66: gb\_vil9:\*  
67: gb\_vil10:\*  
68: gb\_vil11:\*  
69: gb\_vil12:\*  
70: gb\_vil13:\*  
71: gb\_vil14:\*  
72: gb\_vil15:\*  
73: gb\_vil16:\*  
74: gb\_vil17:\*  
75: gb\_vil18:\*  
76: gb\_vil19:\*  
77: gb\_vil20:\*  
78: gb\_vil21:\*  
79: gb\_vil22:\*  
80: gb\_vil23:\*  
81: gb\_vil24:\*  
82: gb\_vil25:\*  
83: gb\_vil26:\*  
84: gb\_vil27:\*  
85: gb\_vil28:\*  
86: gb\_vil29:\*  
87: gb\_vil30:\*  
88: gb\_vil31:\*  
89: gb\_vil32:\*  
90: gb\_vil33:\*  
91: gb\_vil34:\*  
92: gb\_vil35:\*  
93: gb\_vil36:\*  
94: gb\_vil37:\*  
95: gb\_vil38:\*  
96: gb\_vil39:\*  
97: gb\_vil40:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2199	95.6	2323	89	AF326739 Homo sapi
2	2128	92.5	2329	9	AX056687 Sequence
3	2061	89.6	2296	89	AF331844 Homo sapi
4	2036	88.5	21501	89	AF326736 Homo sapi
5	2036	88.5	94752	85	AC003098 Homo sapi
6	1915	83.2	151780	72	AC055813 Homo sapi
7	661	28.7	177744	75	AC073954 Homo sapi
8	201	8.7	177744	75	AC073954 Homo sapi

## ALIGNMENTS

JOURNAL	Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
	220th St. SE, Bothell, WA 98021, USA
FEATURES	Location/Qualifiers
source	1..2323
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="17"
	/map="17q21"
5'/UTR	1..47
CDS	48..689
	/codon_start=1
	/product="sclerostin"
	/protein_id="AAK13454.1"
	/db_xref="GI:13161020"
	/translation="MQLPLALCIWLIVHAPFVYVGOGMOAFKNDATETIIPELGEYF
	EPPELEENKNTKMNARENGRPHHPETTKVSYSCRELHAFRYVLDGCSARKVPEF
	LVGSCGCEPARLDPAIIGRKMRPSPDRCLPDRI RQRIQLLCPGEAPRAKRVV"
	LVASCKRLTRFHNHOSLEKDFTEARLPQKRRPRAKSAKMAQLENNY"
	690..2323
BASE COUNT	590 a 635 c 614 g 484 t
ORIGIN	
Query Match	95.6%; Score 2199; DB 89; Length 2323;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2299; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 agaacctgtctactcttgaaagctgtgccttcctctctgtgctgtgtacatgcagctccac 60
DB	1 AGAAGCTGTCTACTCTGGAAGTGGCGTCCCTCTTGCGTGTACCATGCACCTCCAC 60
QY	61 tggacctgtgtctcatctgtcctgtctgtgtacacacagccttcctgttagtggaggccagg 120
DB	61 TGGCCCTGTGTCTCTGTCTCTGCTGTGTACACAGCCTTCCGTGTGTAGTGAAGGCCAGG 120
QY	121 ggttgacagagcttcaagaatgtagtcacacaggaatcatctgcgagctcggagagtlaccgcg 180
DB	121 GGTGGCAGGGGTTCAGAGATGATGTCACGAGAAATCATCTCCGAGCTGGAGAGTACCCTGG 180
QY	181 agcctccacccgagctcggagaaacaacaagacatlgaaacccggcgagagaaaggagcgagc 240
DB	181 AGCCTCACCGGAGCTGGAGAACACAAAGACCATGAACCGGCGGAGAACGGAGGCGCG 240
QY	241 ctccccacacccccctttgagaccaaagacggtgtccagtagaagctgtgcgagctgtgact 300
DB	241 CTCCTCCACCACTCCCTTTGAGACCAAGACGTTCCGAGTATACGTTGCGGAGACTGCAC 300
QY	301 tcaaccgctactcgtgacacgga tgggcgctgcgcgagcgacgacacggttaccgagctgtgtgt 360
DB	301 TCACCGGCTACTGTCGACGATGGGCGGTGCGCGAGCGCCAAACGCGGTGACCGAGCTGTGT 360
QY	361 gcttcgagcagctgtgcgcccgcgcgcgtgtctgtgtcccaacgcca tgcgcgcgcgaagtgtgt 420
DB	361 GCTTCGGCCAGTTCGGCGCCGCGCGCTGTGTCCCAACGCCATGAGGCGCGGCAAGTGTGT 420
QY	421 ggcgacctaa tgggcgcgagcttcgcgtctga tccccgagacgcgtatccgcgcgcgcgcgtgc 480
DB	421 GCGGACCTAATGGGCCCCGACTTCGCTGTGATCCCGAGCGCTACCGGCGCGCAGCGGTGC 480
QY	481 agctgtctgttcccggtgtgtgagcgcgcgcgcgaggtgtgcctcgtgtgtgtgtgtgtgt 540
DB	481 AGCTGTGTGTCCCGGTGTGTGAGCGCGCGCGCGCGCAGCAAGTGTGCTGCTGTGTGTGTGT 540
QY	541 gcaagtgcaagcgctccacccgcttccacaaacacagtcggagcttcaagaagatttggagacg 600
DB	541 GCAAGTCAAGCGCTCCACCGCTTCCACAAACAGTCCGAGCTCAAGGACTTGGGAGCGG 600
QY	601 aggcgcctcgcgcgcgcgaagagcggaagacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 660
DB	601 AGGCGCGTCCGCGCGCAGAGAGGCGCGGAAGCGCGCGCGCGCGCGCGGAGCGCAAGGCCA 660
QY	661 accaagcgagctgtgagaaacgcttactatagagccgcgcgcgccttcccaaccgcgcgggc 720
DB	661 ACGGCGGAGCTGTGAGAAACGCTTACTATAGAGCCGCGCGCGCGCGCGCGCGCGCGCGCG 720



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2001, 00:15:05 ; Search time 272.77 Seconds  
(without alignments)  
5296.776 Million cell updates/sec

Title: US-09-668-021-5  
Perfect score: 2301  
Sequence: 1 agagcctgtcactgtaag.....caatgatcatgaccgaag 2301

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapept 60.0

Database: 730101 seqs, 313950809 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

N\_Geneseq\_0601:\*

1: /SID58/gcgdata/geneseq/geneseqn/NA1980.DAT:\*

2: /SID58/gcgdata/geneseq/geneseqn/NA1981.DAT:\*

3: /SID58/gcgdata/geneseq/geneseqn/NA1982.DAT:\*

4: /SID58/gcgdata/geneseq/geneseqn/NA1983.DAT:\*

5: /SID58/gcgdata/geneseq/geneseqn/NA1984.DAT:\*

6: /SID58/gcgdata/geneseq/geneseqn/NA1985.DAT:\*

7: /SID58/gcgdata/geneseq/geneseqn/NA1986.DAT:\*

8: /SID58/gcgdata/geneseq/geneseqn/NA1987.DAT:\*

9: /SID58/gcgdata/geneseq/geneseqn/NA1988.DAT:\*

10: /SID58/gcgdata/geneseq/geneseqn/NA1989.DAT:\*

11: /SID58/gcgdata/geneseq/geneseqn/NA1990.DAT:\*

12: /SID58/gcgdata/geneseq/geneseqn/NA1991.DAT:\*

13: /SID58/gcgdata/geneseq/geneseqn/NA1992.DAT:\*

14: /SID58/gcgdata/geneseq/geneseqn/NA1993.DAT:\*

15: /SID58/gcgdata/geneseq/geneseqn/NA1994.DAT:\*

16: /SID58/gcgdata/geneseq/geneseqn/NA1995.DAT:\*

17: /SID58/gcgdata/geneseq/geneseqn/NA1996.DAT:\*

18: /SID58/gcgdata/geneseq/geneseqn/NA1997.DAT:\*

19: /SID58/gcgdata/geneseq/geneseqn/NA1998.DAT:\*

20: /SID58/gcgdata/geneseq/geneseqn/NA1999.DAT:\*

21: /SID58/gcgdata/geneseq/geneseqn/NA2000.DAT:\*

22: /SID58/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2301	100.0	2301	21	AAA29056 Human TGF-beta bin
2	2250	97.8	2301	21	AAA29062 Human TGF-beta bin
3	2199	95.6	2301	21	AAA29055 Human TGF-beta bin
4	2148	93.4	2301	21	AAA29061 Mutant human TGF-b
5	2128	92.5	2329	22	AAA91023 Human secreted pro
6	2036	88.5	5680	21	AAA94049 Human TGF-beta bin
7	2036	88.5	9301	21	AAA29064 Human TGF-beta bin
8	563	24.5	642	21	AAA94051 Human DAN/Cerberus
9	422	18.3	1104	21	AAA94050 Human DAN/Cerberus
10	108	4.7	196	21	AAAC05741 Human secreted pro
11	108	4.7	198	20	AAAX0842 Human secreted pro

12	92	4.0	642	21	AAA29057	Vervet TGF-beta bi
13	83	3.6	532	21	AAA29060	Bovine TGF-beta bi
14	69	3.0	638	21	AAA29063	Murine TGF-beta bi
15	69	3.0	35828	21	AAA29063	Murine TGF-beta bi
16	68	3.0	674	21	AAA29059	Rat TGF-beta bindi
17	40	1.7	40	22	AAA91036	Probe for Human se
18	39	1.2	27	22	AAA94042	Human hDCR6 exon 1
19	28	1.2	27	22	AAA94042	PCR primer for Hum
20	27	1.2	38	21	AAA94048	Human hDCR6 exon 4
21	27	1.2	1303	21	AAAC99909	Human secreted pro
22	26	1.1	30	21	AAA94047	Human hDCR6 exon 4
23	26	1.1	49	13	AAO33591	Microsatellite seq
24	26	1.1	320	20	AAAX57405	Rat U3 gene trap d
25	26	1.1	355	21	AAA31528	Plant microsatelli
26	26	1.1	421	21	AAA80642	Human secreted pro
27	26	1.1	559	21	AAAC58916	Eucalyptus grandis
28	26	1.1	597	16	AAO95193	Simple tandem repe
29	26	1.1	893	16	AAO92529	P. communis (pear)
30	26	1.1	1568	21	AAAC37436	Arabidopsis thalia
31	26	1.1	2295	21	AAA80610	Human CD33-like se
32	26	1.1	2483	17	AAAT45982	Human cocaine and
33	26	1.1	2483	19	AAAV57895	Human cocaine and
34	26	1.1	2483	20	AAZ25666	Human cocaine and
35	26	1.1	2764	20	AAZ34109	Human PRO940 nucle
36	26	1.1	2764	21	AAAC78510	Human PRO940 (UNQ4
37	26	1.1	2919	18	AAAT51124	Human homeoprotein
38	26	1.1	3099	20	AAV99912	Human sialoadhesin
39	26	1.1	3099	20	AAV99911	Human sialoadhesin
40	26	1.1	3515	22	AAAF75343	Human TGF-beta rec
41	26	1.1	4011	20	AAAK60124	DNA sequence of th
42	26	1.1	4315	21	AAA60620	Mouse knox1 promot
43	26	1.1	6083	21	AAA51398	Chromosome 16q tum
44	26	1.1	7715	19	AAV43042	Mus musculus Cgamm
45	26	1.1	8760	15	AAO73473	Porcine pro-interl
46	26	1.1	9840	22	AAAC85482	Murine neuropeptid
47	26	1.1	16442	18	AAAX83006	Partial mouse WRN
48	26	1.1	23241	22	AAAF97870	Human neuroblastom
49	26	1.1	23241	22	AAAF97871	Human neuroblastom
50	26	1.1	49999	20	AAZ23895	Murine LOBO homolo

#### ALIGNMENTS

RESULT	1
AAA29056	AAA29056 standard; cDNA; 2301 BP.
XX	XX
AC	AAA29056:
XX	XX
DT	12-SEP-2000 (first entry)
XX	XX
DE	Human TGF-beta binding protein (BEER) variant V101 cDNA.
XX	XX
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW	BEER; variant; V101; gene therapy; antisense therapy; fracture;
KW	Chromosome 17q12-21; bone mineralization; ss.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	48..689
FT	/tag= a
FT	/label= BEER_variant_V101
FT	/product= TGF-beta_binding_protein
PN	MO2000032773-A1.
XX	XX
PD	08-JUN-2000.
XX	XX
PF	24-NOV-1999.
XX	XX
PR	27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.  
 PA  
 XX  
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
 PI Van Ness J, Winkler DG;  
 XX  
 XX WPI: 2000-412321/35.  
 DR P-PSDB; AAY96430.  
 DR  
 XX  
 XX Nucleic acids (I) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 XX  
 XX Claim 1: Page 118-119; 162pp; English.

CC This CDNA encodes a variant human transforming growth factor-beta  
CC (TGF-beta) binding protein designated BEER V101. The encoded protein  
CC comprises a substitution of Isoleucine for the wild-type valine at  
CC residue 10. The CDNA and protein may be used for prevention, treatment  
CC and diagnosis of diseases associated with inappropriate BEER expression.  
CC For example, they may be used to treat disorders associated with  
CC decreased TGF-beta BP expression. The CDNA or vectors may be administered  
CC to treat diseases by rectifying mutations or deletions in a patient's  
CC genome that affect the activity of BEER by expressing inactive proteins  
CC or to supplement the patients own production of BEER polypeptides. The  
CC nucleic acids may be used for recombinant production of BEER, gene  
CC therapy, antisense therapy, as probes for diagnostic assays and for  
CC functional studies. BEER may be used to raise antibodies and for  
CC identification of BEER modulators. BEER antagonists may be used to  
CC increase bone mineral content for the treatment of disorders such as  
CC osteopenia, osteoporosis, fractures and other disorders associated with  
CC low mineral content.

XX Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;

XX

Query Match	100.0%	Score 2301:	DB 21:	Length 2301:
Best Local Similarity	100.0%	Pred. No. 0:		
Matches 2301: Conservative	0:	Mismatches	0:	Indels
			0:	Gaps
				0:
QY	1	agagcctgtgtctactcggagagctgagcgtgtgcctccctctctgtgtatccacgacgtccacc	60	
Db	1	agagcctgtgtctactcggagagctgagcgtgtgcctccctctctgtgtatccacgacgtccacc	60	
QY	61	tggccctgtgtctatctctgcctctcgtgtacaaacagccttcctgtatgtgagggccag	120	
Db	61	tggccctgtgtctatctctgcctctcgtgtacaaacagccttcctgtatgtgagggccag	120	
QY	121	ggctgcagcgctgtcacaagatgaatgcacaggaatatccgcagctcgcagaggtacc	180	
Db	121	ggctgcagcgctgtcacaagatgaatgcacaggaatatccgcagctcgcagaggtacc	180	
QY	181	agcctccacccgagactgtgagaaacaacagaccatgaaccggcgaggaacgagggcgc	240	
Db	181	agcctccacccgagactgtgagaaacaacagaccatgaaccggcgaggaacgagggcgc	240	
QY	241	ctcccccacacccctctgagaccaaagacgtgtccgagatcacgcttgcgcgcgtgcact	300	
Db	241	ctcccccacacccctctgagaccaaagacgtgtccgagatcacgcttgcgcgcgtgcact	300	
QY	301	tcaaccgctacgttgacgcgatalgagccgttgcgcgagagcccaagccggttcaacagctgtgt	360	
Db	301	tcaaccgctacgttgacgcgatalgagccgttgcgcgagagcccaagccggttcaacagctgtgt	360	
QY	361	gtctcgcgcagctgagccgcgcgcgtctgtctgtcccaacgcacatcgccgcgcgaagtgt	420	
Db	361	gtctcgcgcagctgagccgcgcgcgtctgtctgtcccaacgcacatcgccgcgcgaagtgt	420	
QY	421	ggcgcgcctagctgtggccgcgaacttcgcctgtatactccgcacgcctacacgcgcgcgtgc	480	
Db	421	ggcgcgcctagctgtggccgcgaacttcgcctgtatactccgcacgcctacacgcgcgcgtgc	480	
QY	481	agcctcgtgtgtcccggtgtgtgagcgccgcgcgcgcgcgaagtgtgcctctgtgtgcctcgt	540	

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 21:18:58 ; Search time 3376.69 Seconds  
(without alignments)  
6441.520 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301

Sequence: 1 agagcctgctcactcgaag.....caatgatcatgcacgaaag 2301

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Search: 10228115 seqs, 4726426750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

EST:\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*  
10: gb\_est10.\*  
11: gb\_est11.\*  
12: gb\_est12.\*  
13: gb\_est13.\*  
14: gb\_est14.\*  
15: gb\_est15.\*  
16: gb\_est16.\*  
17: gb\_est17.\*  
18: gb\_est18.\*  
19: gb\_est19.\*  
20: gb\_est20.\*  
21: gb\_est21.\*  
22: gb\_est22.\*  
23: gb\_est23.\*  
24: gb\_est24.\*  
25: gb\_est25.\*  
26: gb\_est26.\*  
27: gb\_est27.\*  
28: gb\_est28.\*  
29: gb\_est29.\*  
30: gb\_est30.\*  
31: gb\_est31.\*  
32: gb\_est32.\*  
33: gb\_est33.\*  
34: gb\_est34.\*  
35: gb\_est35.\*  
36: gb\_est36.\*  
37: gb\_est37.\*  
38: gb\_est38.\*  
39: gb\_est39.\*  
40: gb\_est40.\*  
41: gb\_est41.\*  
42: gb\_est42.\*  
43: gb\_est43.\*

44: em\_esthum10.\*  
45: em\_esthum11.\*  
46: em\_esthum12.\*  
47: em\_esthum13.\*  
48: em\_esthum14.\*  
49: em\_esthum15.\*  
50: em\_esthum16.\*  
51: em\_esthum17.\*  
52: em\_esthum18.\*  
53: em\_esthum19.\*  
54: em\_esthum20.\*  
55: em\_esthum21.\*  
56: em\_esthum22.\*  
57: em\_esthum23.\*  
58: em\_esthum24.\*  
59: em\_esthum25.\*  
60: em\_esthum26.\*  
61: em\_esthum27.\*  
62: em\_esthum28.\*  
63: em\_esthum29.\*  
64: em\_esthum30.\*  
65: em\_esthum31.\*  
66: em\_esthum32.\*  
67: em\_esthum33.\*  
68: em\_esthum34.\*  
69: em\_esthum35.\*  
70: em\_esthum36.\*  
71: em\_esthum37.\*  
72: em\_esthum38.\*  
73: em\_esthum39.\*  
74: em\_esthum40.\*  
75: em\_esthum41.\*  
76: em\_esthum42.\*  
77: em\_esthum43.\*  
78: em\_esthum44.\*  
79: em\_esthum45.\*  
80: em\_esthum46.\*  
81: em\_esthum47.\*  
82: em\_esthum48.\*  
83: em\_esthum49.\*  
84: em\_esthum50.\*  
85: em\_esthum51.\*  
86: em\_esthum52.\*  
87: em\_esthum53.\*  
88: em\_esthum54.\*  
89: em\_esthum55.\*  
90: em\_esthum56.\*  
91: em\_esthum57.\*  
92: em\_esthum58.\*  
93: em\_esthum59.\*  
94: em\_esthum60.\*  
95: em\_esthum61.\*  
96: em\_esthum62.\*  
97: em\_esthum63.\*  
98: em\_esthum64.\*  
99: em\_esthum65.\*  
100: em\_esthum66.\*  
101: em\_esthum67.\*  
102: em\_esthum68.\*  
103: em\_esthum69.\*  
104: em\_esthum70.\*  
105: em\_esthum71.\*  
106: em\_esthum72.\*  
107: em\_esthum73.\*  
108: em\_esthum74.\*  
109: em\_esthum75.\*  
110: em\_esthum76.\*  
111: em\_esthum77.\*  
112: em\_esthum78.\*  
113: em\_esthum79.\*  
114: em\_esthum80.\*  
115: em\_esthum81.\*  
116: em\_esthum82.\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est110:\*  
191: gb\_est111:\*  
192: gb\_est112:\*  
193: gb\_est113:\*  
194: gb\_est114:\*  
195: gb\_est115:\*  
196: gb\_est116:\*  
197: gb\_est117:\*  
198: gb\_est118:\*  
199: gb\_est119:\*  
200: gb\_est120:\*  
201: gb\_est121:\*  
202: gb\_est122:\*  
203: gb\_est123:\*  
204: gb\_est124:\*  
205: gb\_est125:\*  
206: gb\_est126:\*  
207: gb\_est127:\*  
208: gb\_est128:\*  
209: gb\_est129:\*  
210: gb\_est130:\*  
211: gb\_est131:\*  
212: gb\_est132:\*  
213: gb\_est133:\*  
214: gb\_est134:\*  
215: gb\_est135:\*  
216: gb\_est136:\*  
217: gb\_est137:\*  
218: gb\_est138:\*  
219: gb\_est139:\*  
220: gb\_est140:\*  
221: gb\_est141:\*  
222: gb\_est142:\*  
223: gb\_est143:\*  
224: gb\_est144:\*  
225: gb\_est145:\*  
226: gb\_est146:\*  
227: gb\_est147:\*  
228: gb\_est148:\*  
229: gb\_est149:\*  
230: gb\_est150:\*  
231: gb\_est151:\*  
232: gb\_est152:\*  
233: gb\_est153:\*  
234: gb\_est154:\*  
235: gb\_est155:\*  
236: gb\_est156:\*  
237: gb\_est157:\*  
238: gb\_est158:\*  
239: gb\_est159:\*  
240: gb\_est160:\*  
241: gb\_est161:\*  
242: gb\_est162:\*  
243: gb\_est163:\*  
244: gb\_est164:\*  
245: gb\_est165:\*  
246: gb\_est166:\*  
247: gb\_est167:\*  
248: gb\_est168:\*  
249: gb\_est169:\*  
250: gb\_est170:\*  
251: gb\_est171:\*  
252: gb\_est172:\*  
253: gb\_est173:\*  
254: gb\_est174:\*  
255: gb\_est175:\*  
256: gb\_est176:\*  
257: gb\_est177:\*  
258: gb\_est178:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepel BW,  
PI Van Ness J, Winkler DG;

PI Van Ness J, Winkler DG;

DR WPT; 2000-412321/35.

DR P-PSDB; AAY96429.

PT Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures

PT osteoporosis and fractures

PT osteoporosis and fractures

PS Claim 1; Page 114-115; 162pp; English

CC This CD encodes a human transforming growth factor-beta (TGF-beta)  
CC binding protein designated BEER. The hBEER gene has been localized  
CC to the chromosome 17q12-21. The CDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The CDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patients own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other  
CC disorders associated with low mineral content.



Query Match 1.1%; Score 26; DB 2; Length 597;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 gagagagagagagagagagagaga 1353  
|||||  
Db 399 GAGAGAGAGAGAGAGAGAGAGA 424

RESULT 2  
US-08-276-452A-66/C  
; Sequence 66, Application US/08276452A  
; Patent No. 5646029  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Chao-Guang  
; APPLICANT: Mau, Shao-Lim  
; APPLICANT: Du, He  
; APPLICANT: Gane, Allison M  
; APPLICANT: Bacic, Antony  
; APPLICANT: Clarke, Adrienne E  
; TITLE OF INVENTION: Plant Arabidnagalactan Protein (AGP) Genes  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: United States of America  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/276,452A  
; FILING DATE: 18-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Caruthers, Jennie M.  
; REGISTRATION NUMBER: 34,464  
; REFERENCE/DOCKET NUMBER: 27-91A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303)499-8080  
; TELEFAX: (303)499-8089  
; TELEX: 49617824  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 893 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 70..504  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 70..138 /note= "Putative secretion signal"  
; OTHER INFORMATION: peptide"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 70..339  
; OTHER INFORMATION: /note= "Amino acids 70-138, 38-53,  
; and 71-90 are sequences which match the peptide  
; OTHER INFORMATION: sequences obtained by protein sequencing"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 87..324  
; OTHER INFORMATION: /note= "Amino acids 40, 44, 45, 49,  
; 50, 74, 76, 80, 81, and 85 are hydroxylated"

OTHER INFORMATION: prolines"  
US-08-276-452A-66  
Query Match 1.1%; Score 26; DB 1; Length 893;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 gagagagagagagagagagagaga 1353  
|||||  
Db 64 GAGAGAGAGAGAGAGAGAGAGA 39

RESULT 3  
US-08-798-744-66/C  
; Sequence 66, Application US/08798744  
; Patent No. 5830747  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Chao-Guang  
; APPLICANT: Mau, Shao-Lim  
; APPLICANT: Du, He  
; APPLICANT: Gane, Allison M  
; APPLICANT: Bacic, Antony  
; APPLICANT: Clarke, Adrienne E  
; TITLE OF INVENTION: Plant Arabidnagalactan Protein (AGP) Genes  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: United States of America  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/798,744  
; FILING DATE: 13-FEB-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/276,452  
; FILING DATE: 18-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Caruthers, Jennie M.  
; REGISTRATION NUMBER: 34,464  
; REFERENCE/DOCKET NUMBER: 27-91A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303)499-8080  
; TELEFAX: (303)499-8089  
; TELEX: 49617824  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 893 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 70..504  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 70..138 /note= "Putative secretion signal"  
; OTHER INFORMATION: peptide"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 70..339  
; OTHER INFORMATION: /note= "Amino acids 70-138, 38-53,  
; and 71-90 are sequences which match the peptide"



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 22:25:30 ; Search time 3376.69 Seconds

(without alignments)  
6441.520 Million cell updates/sec

Title: US-09-668-021-5

Perfect score: 2301

Sequence: 1 agagcctgtgctactggaag.....caatgaatcatgacgaaag 2301

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Word size : 0  
10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Listing first 50 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: em_est100:*
170: em_est101:*
171: em_est102:*
172: em_est103:*
173: em_est104:*
174: em_est105:*
175: em_est106:*
176: em_est107:*
177: em_est108:*
178: em_est109:*
179: em_est110:*
180: em_est111:*
181: em_est112:*
182: em_est113:*
183: em_est114:*
184: em_est115:*
185: em_est116:*
186: em_est117:*
187: em_est118:*
188: em_est119:*
189: em_est120:*
```

```
190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:32:16 ; Search time 54.37 seconds

(without alignments)  
298.422 Million cell updates/sec

Title: US-09-668-021-2

Perfect score: 213

Sequence: 1 MOLPLALCLVLLVHTAFRV.....KPRPRASAKANQAELENAY 213

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : PIR\_68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	428	1	transforming prote
2	8	3.8	429	2	elk1 protein - mou
3	7	3.3	130	2	hypothetical prote
4	7	3.3	148	2	hypothetical prote
5	7	3.3	183	2	H-2 class II histo
6	7	3.3	197	2	hypothetical prote
7	7	3.3	197	2	hypothetical prote
8	7	3.3	305	2	yabc protein homol
9	7	3.3	312	2	conserved hypothet
10	7	3.3	321	2	conserved hypothet
11	7	3.3	328	2	hypothetical prote
12	7	3.3	398	1	polyferredoxin 6x2
13	7	3.3	408	2	fimbrial assembly
14	7	3.3	412	1	cathepsin D (EC 3.
15	7	3.3	510	2	legumin 31 precurs
16	7	3.3	512	1	inclusion body mat
17	7	3.3	562	2	hypothetical prote
18	7	3.3	587	2	hypothetical prote
19	7	3.3	602	2	NADH dehydrogenase
20	7	3.3	632	1	surface glycoprote
21	7	3.3	681	2	outer membrane hem
22	7	3.3	781	1	protein kinase Dra
23	7	3.3	917	1	hexokinase (EC 2.7
24	7	3.3	946	2	hypothetical prote
25	7	3.3	1940	2	myosin heavy chain
26	7	3.3	6420	2	polyketide synthas
27	6	2.8	49	2	neurotoxin rx2 - s
28	6	2.8	65	2	malx protein - Kle
29	6	2.8	83	2	tightly associated

## ALIGNMENTS

RESULT 1

TVHUEK

transforming protein elk-1 - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999

C:Accession: A41354; S54721

R:Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.

Science 244, 66-70, 1989

A:Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near transloc

A:Reference number: A41354; MUID:89203250

A:Accession: A41354

A:Molecule type: mRNA

A:Residues: 1-428 <RAO>

A:Cross-references: GB:M25269; NID:g538208; PIDN:AAA52384.1; PID:g538209

R:Gille, H.; Kortgenjann, M.; Thomae, O.; Moomaw, C.; Slaughter, C.; Cobb, M.H.; Shaw,

EMBO J. 14, 951-962, 1995

A:Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and

A:Reference number: S54721; MUID:95196756

A:Accession: S54721

A:Status: preliminary

A:Molecule type: protein

A:Residues: 318-328, 'XX', 331, 336-364, 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X'

C:Genetics:

A:Gene: GDB:ELK1

A:Cross-references: GDB:119867; OMIM:311040

A:Map position: Xp11.2-Xp11.2

C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology

C:Keywords: DNA binding; oncogene; transforming protein

F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match

Best Local Similarity 3.8%; Score 8; DB 1; Length 428;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 PQGRKPR 196

|||||

Db 310 PQGRKPR 317

RESULT 2

JC4965

elk1 protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 16-Jul-1999

C:Accession: JC4965; I48339; I48340; S54908

R:Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatannens, B.; Begue, A.; Stehelin, D.

Gene 174, 185-188, 1996  
 A:Title: Structure and organization of the mouse elk1 gene.

A:Reference number: JC4965; MUID:97017146

A:Accession: JC4965

A:Molecule type: mRNA

A:Residues: 1-429 <GRE>

A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635

A:Experimental source: embryo

R:Giovane, A.; Pintzas, A.; Maira, S.M.; Sobieszczuk, P.; Wasyluk, B.

Genes Dev. 8, 1502-1513, 1994

A:Title: Net, a new ets transcription factor that is activated by Ras.

A:Reference number: A53837; MUID:95047310

A:Accession: I48339

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-429 <RES>

A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635

A:Accession: I48340

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 5-132, 'T', 134-224 <RE2>

A:Cross-references: EMBL:Z36939; NID:g535922; PIDN:CAA85391.1; PID:g535923

C:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which b

C:Genetics:

A:Gene: elk1

A:Introns: 70/3; 219/3; 363/3; 397/3

C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology

F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 3.3%; Score 8; DB 2; Length 429;

Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 POKGRKPR 196

|||||||

Db 311 POKGRKPR 318

RESULT 3

hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: D83305

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337

A:Accession: D83305

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-130 <STO>

A:Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AAG06110.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2722

Query Match 3.3%; Score 7; DB 2; Length 130;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 KVLRLVAS 164

|||||||

Db 12 KVLRLVAS 18

RESULT 4

hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)

C83091

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: C83091

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: C83091

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-148 <STO>

A:Cross-references: GB:AE004858; GB:AE004091; NID:g9950668; PIDN:AAG07829.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA4441

Query Match 3.3%; Score 7; DB 2; Length 148;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 112 ARLLPNA 118

|||||||

Db 26 ARLLPNA 32

RESULT 5

B37410

H-2 class II histocompatibility antigen A-2 beta chain - spiny mouse (Mus saxicola) (

C:Species: Mus saxicola

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 21-Jan-2000

C:Accession: B37410

R:Cam, P.; Jouvin-Marche, E.; LeGuern, C.; Marche, P.N.

Eur. J. Immunol. 20, 1337-1343, 1990

A:Title: Structure of class II genes in wild mouse Mus saxicola: functional and evolu

A:Reference number: A37410; MUID:90316177

A:Accession: B37410

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-183 <CAM>

A:Cross-references: GB:M30158

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

F:105-170/Domain: immunoglobulin homology <IMM>

Query Match 3.3%; Score 7; DB 2; Length 183;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 EIIPELG 41

|||||||

Db 157 EIIPELG 163

RESULT 6

G81057

hypothetical protein NMB1656 [imported] - Neisseria meningitidis (strain MC58 serogro

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001

C:Accession: G81057

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: G81057

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-197 <TET>

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:35:31 : Search time 31.03 Seconds  
(without alignments)  
235.141 Million cell updates/sec

Title: US-09-668-021-2

Perfect score: 213

Sequence: 1 MQLPLALVLVCLVHTAFRV.....KPRPRRSKANKQALENAY 213

Scoring table:

OLIGO

Gapop 60.0 , Gapept 60.0

Archived: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	428	1	ELK1_HUMAN
2	8	3.8	429	1	ELK1_MOUSE
3	7	3.3	110	1	RL22_LEPIN
4	7	3.3	305	1	YB44_SYNY3
5	7	3.3	412	1	CATD_HUMAN
6	7	3.3	512	1	IBMP_FMYD
7	7	3.3	602	1	N05M_DTDMA
8	7	3.3	632	1	VGLG_SYNY
9	7	3.3	781	1	KRAF_DROME
10	7	3.3	917	1	HXK2_MOUSE
11	7	3.3	917	1	HXK2_MOUSE
12	6	2.8	49	1	TX25_PPHNT
13	6	2.8	65	1	MALX_KLEPN
14	6	2.8	97	1	OSTC_CHICK
15	6	2.8	97	1	VE7_HPVA4
16	6	2.8	101	1	APC2_MACFA
17	6	2.8	110	1	RL22_ECOLI
18	6	2.8	111	1	RL22_ACHLA
19	6	2.8	112	1	RL22_SPICI
20	6	2.8	113	1	RL22_THTH
21	6	2.8	115	1	TLAF_HUMAN
22	6	2.8	115	1	TLAF_MOUSE
23	6	2.8	118	1	REV_HVILW
24	6	2.8	123	1	RS12_ECOLI
25	6	2.8	131	1	RL22_PHSI
26	6	2.8	133	1	TL4_BOVIN
27	6	2.8	135	1	TL4_CAPI
28	6	2.8	135	1	TL4_SHEEP
29	6	2.8	146	1	YIIR_ECOLI
30	6	2.8	150	1	R19E_PYROO
31	6	2.8	167	1	B3AR_MERUN
32	6	2.8	168	1	TRIC_CHICK
33	6	2.8	175	1	RL6A_YEAST

## ALIGNMENTS

RESULT	ID	ELK1_HUMAN	STANDARD;	PRT;	428 AA.
AC	P19419	075606; Q9UJW4; 095058;			
DT	01-NOV-1990	(Rel. 16, Created)			
DT	01-OCT-2000	(Rel. 40, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	ETS-DOMAIN PROTEIN ELK-1.				
OS	ELK1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=89203250; PubMed=2539641;				
RA	Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,				
RA	Reddy E.S.P.,				
RT	"elk, tissue-specific ets-related genes on chromosomes X and 14 near				
RL	translocation breakpoints.";				
RL	Science 244:66-70(1989).				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=99013876; PubMed=9795224;				
RA	Hardinath N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.,				
RT	"The human elk-1 gene family: the functional gene and two processed				
RL	pseudogenes embedded in the 19h locus.";				
RL	Gene 221:215-224(1998).				
RP	SEQUENCE FROM N.A.				
RA	Grafham D.,				
RL	Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RA	Arvey D.N.T., Kovar H.,				
RT	"Novel family members HUER71, ELFR, and ELKV among ETS-related genes				
RT	coexpressed with EMS-FLI1 in Ewing tumor cell lines.";				
RL	Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.				
RP	DOMAINS.				
RP	MEDLINE=92334979; PubMed=1630903;				
RA	Janknecht R., Nordheim A.,				
RT	"Elk-1 protein domains required for direct and SRF-assisted				
RT	DNA-binding.";				
RL	Nucleic Acids Res. 20:3317-3324(1992).				
CC	-1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA				
CC	SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE				
CC	FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE				
CC	ELEMENT.				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.				
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE), AND 2/ELKV; ARE				

34	6	2.8	182	1	YCYO_YEAST	P25654 saccharomyc
35	6	2.8	196	1	RETB_CHICK	P41263 gallus gall
36	6	2.8	206	1	BREF1_HUMAN	P20290 homo sapien
37	6	2.8	210	1	TRPF_KLUIA	P13997 kuyetomyc
38	6	2.8	213	1	PYRE_HAETN	P43855 haemophilus
39	6	2.8	216	1	PGFH_HUMAN	O60258 homo sapien
40	6	2.8	216	1	PGFH_MOUSE	O70627 mus musculu
41	6	2.8	230	1	219_HUMAN	P98173 homo sapien
42	6	2.8	234	1	TNFR_HUMAN	P32971 homo sapien
43	6	2.8	237	1	NRL_MOUSE	P54846 mus musculu
44	6	2.8	247	1	PS12_XENLA	O99vq1 xenopus lae
45	6	2.8	248	1	PS17_XENLA	O99vq6 xenopus lae
46	6	2.8	248	1	PSA7_HUMAN	O14818 homo sapien
47	6	2.8	248	1	PSA7_MOUSE	O922u0 mus musculu
48	6	2.8	249	1	APX1_PEA	P48534 pisum sativ
49	6	2.8	249	1	PSA7_CHICK	O13268 gallus gall
50	6	2.8	254	1	PSA7_RAT	P48004 rattus norv

```

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: LONG AND TESTIS.
CC -1- PTM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,
CC AND ERK).
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25269; AAA52384.1; -.
DR EMBL: AF080616; AAC82466.1; -.
DR EMBL: AL009172; CA15659.1; -.
DR EMBL: AF000672; AAD00862.1; -.
DR PIR: A41354; TVHUEK.
DR HSSP: P14921; 2STW.
DR TRANSFAC: T00250; -.
DR MIM: 311040; -.
DR InterPro: IPR000418; -.
DR Pfam: PF00178; Ets; 1.
DR PRINTS: PR00454; ETSDOMAIN.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation; Alternative splicing.
FT DNA_BIND 5 86 ETS-DOMAIN.
FT VARSPIC 91 95 VAGCS -> SHCAP (IN ISOFORM 2).
FT VARSPIC 96 428 MISSING (IN ISOFORM 2).
FT CONFLICT 183 183 S -> N (IN REF. 1).
SQ SEQUENCE 428 AA; 44888 MW; 68F71F8ADB9D38CA CRC64;

Query Match 3.8%; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKGRKPR 196
DB 310 POKGRKPR 317

RESULT 2
ELK1_MOUSE STANDARD; PRT; 429 AA.
ID ELK1_MOUSE
P41969;
01-NOV-1995 (Rel. 32, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
DE ETS-DOMAIN PROTEIN ELK-1.
GN ELK1.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_Taxid=10090;
OX 11
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL; TISSUE=Embryo;
RX MEDLINE=97017146; PubMed=863747;
RA Grevin D., Ong S., Denhez F., Dehem M., Quatanens B., Begue A.,
RA Stehelin D., Martin P.;
RT "Structure and organization of the mouse elk1 gene.";
RL Gene 174:185-188(1996).
RN 12
RP SEQUENCE OF 5-224 FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=95047310; PubMed=7958835;
RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Maslyk B.;
RT "Net, a new ets transcription factor that is activated by Ras.";
CC

```

```

RL Genes Dev. 8:1502-1513(1994).
CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
CC ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87257; CA60715.1; -.
DR EMBL: Z36939; CA85391.1; -.
DR HSSP: P14921; 2STW.
DR MGD: MGI:101833; ELK1.
DR InterPro: IPR000418; -.
DR Pfam: PF00178; Ets; 1.
DR PRINTS: PR00454; ETSDOMAIN.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation.
FT DNA_BIND 5 86 ETS-DOMAIN.
FT CONFLICT 133 133 P -> T (IN REF. 2).
SQ SEQUENCE 429 AA; 45243 MW; B61B5B97731D54F CRC64;

Query Match 3.8%; Score 8; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKGRKPR 196
DB 311 POKGRKPR 318

RESULT 3
RL22_LEPIN STANDARD; PRT; 110 AA.
ID RL22_LEPIN
AC G9XD31;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L22.
GN RPLV.
OS Leptospira interrogans.
CC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
CC NCBI_Taxid=173;
OX 11
RN SEQUENCE FROM N.A.
RP STRAIN=SEVOVAR LAI;
RX MEDLINE=20088835; PubMed=10620683;
RA Zuerner R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;
RT "Characterization of the Leptospira interrogans S10-spc-alpha
RT operon.";
RL FEMS Microbiol. Lett. 182:303-308(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S rRNA: ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G. L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 9, 2001, 15:34:36 ; Search time 89.8 Seconds  
(without alignments)  
313.819 Million cell updates/sec

Title: US-09-668-021-2

Perfect score: 213  
Sequence: 1 MQLPLALCVLCVLTHTAFRV.....KRRPRARSAKANQALENNAY 213

Scoring table:  
Gapop 60.0 , Gapext 60.0

Archived: 425026 seqs, 132305027 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database:

SPREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rdent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	198	10	Q9FYQ3
2	8	3.8	276	13	Q9YGD6
3	8	3.8	276	13	Q9PT38
4	8	3.8	276	13	Q9PT37
5	8	3.8	276	13	Q9PT36
6	7	3.3	89	10	Q9FH55
7	7	3.3	109	5	Q9GT34
8	7	3.3	114	5	Q9GT34
9	7	3.3	130	2	Q910C0
10	7	3.3	148	2	Q9HVA4
11	7	3.3	183	7	Q31243
12	7	3.3	186	10	Q9SDR7
13	7	3.3	197	2	Q9JTB9
14	7	3.3	197	2	Q9JTB0
15	7	3.3	251	2	Q9JLNO
16	7	3.3	282	10	Q9XFI9
17	7	3.3	312	2	Q9RT99
18	7	3.3	321	2	Q9PQ72
19	7	3.3	398	1	Q00388

20	7	3.3	408	2	Q66951	Q66951 aquifex aeo
21	7	3.3	425	5	Q19348	Q19348 caenorhabdi
22	7	3.3	510	10	Q39521	Q39521 cryptomeria
23	7	3.3	546	13	Q91A13	Q91A13 pimephales
24	7	3.3	562	10	Q81837	Q81837 arabidopsis
25	7	3.3	587	5	Q21432	Q21432 caenorhabdi
26	7	3.3	604	5	Q44003	Q44003 toxoplasma
27	7	3.3	638	4	Q9H1Y7	Q9H1Y7 homo sapien
28	7	3.3	648	4	Q9HA12	Q9HA12 homo sapien
29	7	3.3	667	4	Q9Y2N7	Q9Y2N7 homo sapien
30	7	3.3	681	2	Q9PCB8	Q9PCB8 xylella fas
31	7	3.3	739	5	Q9W4Z3	Q9W4Z3 drosophila
32	7	3.3	782	5	Q9NEH9	Q9NEH9 drosophila
33	7	3.3	802	5	Q96398	Q96398 schistosoma
34	7	3.3	815	4	Q43273	Q43273 homo sapien
35	7	3.3	917	11	Q54892	Q54892 ratu5 norv
36	7	3.3	946	5	Q20143	Q20143 caenorhabdi
37	7	3.3	980	4	Q9NS55	Q9NS55 homo sapien
38	7	3.3	1045	4	Q9UNR9	Q9UNR9 homo sapien
39	7	3.3	1045	4	Q9NS56	Q9NS56 homo sapien
40	7	3.3	1064	4	Q75770	Q75770 homo sapien
41	7	3.3	1064	4	Q9NQ38	Q9NQ38 homo sapien
42	7	3.3	1325	2	Q9W4X3	Q9W4X3 acetobacter
43	7	3.3	1935	5	Q44934	Q44934 loligo peal
44	7	3.3	1940	5	Q02456	Q02456 schistosoma
45	7	3.3	2517	5	Q9W060	Q9W060 drosophila
46	7	3.3	6420	2	P95814	P95814 streptomyce
47	6	2.8	22	2	Q85607	Q85607 streptomyce
48	6	2.8	40	14	Q92266	Q92266 simian hecp
49	6	2.8	47	7	Q31227	Q31227 mus musculu
50	6	2.8	62	10	Q91DG4	Q91DG4 oryza sativ

## ALIGNMENTS

RESULT 1  
ID Q9FYQ3 PRELIMINARY; PRT; 198 AA.  
AC Q9FYQ3;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;  
OC Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, PAC  
clone: P0433F09.\*";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP002539; BAB08182.1; -  
SQ SEQUENCE 198 AA; 21843 MW; 11DD1D90FEAB72P4E CRC64;

Query Match 3.88; Score 8; DB 10; Length 198;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 PGCEAPRA 156  
DB 83 PGCEAPRA 90  
RESULT 2  
QYGD6  
ID QYGD6 PRELIMINARY; PRT; 276 AA.  
AC QYGD6;

```

DR 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DE 01-MAY-2001 (TREMblrel. 16, Last annotation update)
DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99185307; PubMed=10082666;
RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
RT "Cloning and expression of two carbonyl reductase-like 20beta-
hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
trout (Oncorhynchus mykiss).";
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
DR EMBL; AF100930; MAD20217.1; -.
DR HSSP; P50163; 2AE1.
DR InterPro; IPR002198; -.
DR InterPro; IPR002347; -.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDHRDH.
DR PRINTS; PR00080; SDRFAMILY.
KM Oxidoreductase.
SQ SEQUENCE 276 AA; 30111 MW; 135604A3BD2940CD CRC64;

```

```

Query Match          3.8%; Score 8; DB 13; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 28 AFKNDAT 35
Db 93 AFKNDAT 100

```

```

RESULT 3
Q9PT38 PRELIMINARY; PRT; 276 AA.
AC Q9PT38;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99185307; PubMed=10082666;
RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
RT "Cloning and expression of two carbonyl reductase-like 20beta-
hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
trout (Oncorhynchus mykiss).";
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
DR EMBL; AF100933; MAD20992.1; -.
DR HSSP; P50163; 2AE1.
DR InterPro; IPR002198; -.
DR InterPro; IPR002347; -.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDHRDH.
DR PRINTS; PR00080; SDRFAMILY.
KM Oxidoreductase.
SQ SEQUENCE 276 AA; 30239 MW; 134B52054C751C28 CRC64;

```

```

Query Match          3.8%; Score 8; DB 13; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 28 AFKNDAT 35
Db 93 AFKNDAT 100

```

```

RESULT 4
Q9PT37 PRELIMINARY; PRT; 276 AA.
AC Q9PT37;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99185307; PubMed=10082666;
RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
RT "Cloning and expression of two carbonyl reductase-like 20beta-
hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
trout (Oncorhynchus mykiss).";
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
DR EMBL; AF100932; MAD20991.1; -.
DR HSSP; P50163; 2AE1.
DR InterPro; IPR002198; -.
DR InterPro; IPR002347; -.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDHRDH.
DR PRINTS; PR00080; SDRFAMILY.
KM Oxidoreductase.
SQ SEQUENCE 276 AA; 30127 MW; 1350E4211D3422SD CRC64;

```

```

Query Match          3.8%; Score 8; DB 13; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 28 AFKNDAT 35
Db 93 AFKNDAT 100

```

```

RESULT 5
Q9PT36 PRELIMINARY; PRT; 276 AA.
AC Q9PT36;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99185307; PubMed=10082666;
RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
RT "Cloning and expression of two carbonyl reductase-like 20beta-
hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
trout (Oncorhynchus mykiss).";
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:30:31 ; Search time 47.66 Seconds  
(without alignments)  
100.571 Million cell updates/sec

Title: US-09-668-021-2

Perfect score: 213  
Sequence: 1 MQLPLALCLVCLVHTAFRV.....KPRPRARSARAKNAQLEINAY 213

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.2	206	1	US-08-468-847B-2
2	9	4.2	206	1	US-08-468-847B-20
3	9	4.2	206	4	US-09-188-930-159
4	9	4.2	206	4	US-09-188-930-286
5	7	3.3	114	1	US-08-582-257-15
6	7	3.3	114	2	US-08-582-298-15
7	7	3.3	186	4	US-09-475-316A-13
8	7	3.3	412	1	US-08-208-007A-12
9	7	3.3	412	4	US-08-974-691-4
10	7	3.3	917	2	US-08-588-983-16
11	7	3.3	917	3	US-08-588-976-16
12	6	2.8	39	3	US-08-554-385-22
13	6	2.8	68	2	US-08-836-791-5
14	6	2.8	80	1	US-08-011-398B-15
15	6	2.8	80	1	US-08-464-051-15
16	6	2.8	80	2	US-08-462-498-15
17	6	2.8	80	3	US-08-554-385-14
18	6	2.8	88	2	US-08-690-011A-31
19	6	2.8	97	2	US-08-245-511-30
20	6	2.8	97	2	US-08-600-993A-30
21	6	2.8	125	1	US-08-225-989-20
22	6	2.8	125	1	US-08-570-923-20
23	6	2.8	125	4	US-08-580-014-20
24	6	2.8	125	4	US-09-079-785-20
25	6	2.8	163	4	US-09-354-129-12
26	6	2.8	186	1	US-07-960-981-5
27	6	2.8	186	5	PCT-US93-09634-5

28	6	2.8	215	1	US-08-225-989-23	Sequence 23, Appl
29	6	2.8	215	1	US-08-570-923-23	Sequence 23, Appl
30	6	2.8	215	1	US-08-580-014-23	Sequence 23, Appl
31	6	2.8	215	4	US-09-079-785-23	Sequence 23, Appl
32	6	2.8	216	2	US-08-821-637-3	Sequence 3, Appl
33	6	2.8	223	4	US-08-928-941D-22	Sequence 22, Appl
34	6	2.8	223	4	US-09-280-590A-22	Sequence 22, Appl
35	6	2.8	223	1	US-08-225-989-8	Sequence 8, Appl
36	6	2.8	234	1	US-08-570-923-8	Sequence 8, Appl
37	6	2.8	234	1	US-08-580-014-8	Sequence 8, Appl
38	6	2.8	234	4	US-09-079-785-8	Sequence 8, Appl
39	6	2.8	243	2	US-08-319-376-2	Sequence 2, Appl
40	6	2.8	248	2	US-08-701-935-3	Sequence 3, Appl
41	6	2.8	248	3	US-09-134-591-3	Sequence 3, Appl
42	6	2.8	254	2	US-08-701-935-6	Sequence 6, Appl
43	6	2.8	254	3	US-09-134-591-6	Sequence 6, Appl
44	6	2.8	269	2	US-08-727-311-3	Sequence 3, Appl
45	6	2.8	330	4	US-09-188-930-144	Sequence 144, App
46	6	2.8	330	4	US-09-188-930-278	Sequence 278, App
47	6	2.8	343	2	US-08-933-750C-13	Sequence 13, Appl
48	6	2.8	343	2	US-09-234-613-13	Sequence 13, Appl
49	6	2.8	354	2	US-08-700-013B-9	Sequence 9, Appl
50	6	2.8	364	1	US-08-680-726A-56	Sequence 56, Appl

#### ALIGNMENTS

RESULT 1  
US-08-468-847B-2  
; Sequence 2, Application US/08468847B  
; Patent No. 5780263  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
; TITLE OF INVENTION: Human CCM-Like Growth Factor  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468, 847B  
; FILING DATE: 6 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-08-468-847B-2

Query Match 4.2%; Score 9; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 2

US-08-468-847B-20  
; Sequence 20, Application US/08468847B  
; Patent No. 5780263

## GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B

FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J. G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744

INFORMATION FOR SEQ. ID NO: 20:  
SEQUENCE CHARACTERISTICS:

LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID

STRANDEDNESS:  
TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN  
US-08-468-847B-20

## Query Match

Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 3

US-09-188-930-159  
; Sequence 159, Application US/09188930A  
; Patent No. 6150502

## GENERAL INFORMATION:

APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James Greg

TITLE OF INVENTION: Compositions Isolated From Skin Cells  
TITLE OF INVENTION: and Methods For Their Use  
FILE REFERENCE: 11000.1011c1

CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 159  
LENGTH: 206

TYPE: PRT  
ORGANISM: mouse

US-09-188-930-159

Query Match 4.2%; Score 9; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 4

US-09-188-930-286  
; Sequence 286, Application US/09188930A  
; Patent No. 6150502

## GENERAL INFORMATION:

APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene

APPLICANT: Murison, James Greg  
TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: and Methods For Their Use  
FILE REFERENCE: 11000.1011c1

CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 286  
LENGTH: 206

TYPE: PRT  
ORGANISM: Mouse

US-09-188-930-286

Query Match 4.2%; Score 9; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 5

US-08-582-257-15  
; Sequence 15, Application US/08582257  
; Patent No. 5763400

## GENERAL INFORMATION:

APPLICANT: Adams, Michael E.  
APPLICANT: Zltan, Dusan

TITLE OF INVENTION: Ecdysis-Triggering Hormone Compositions  
TITLE OF INVENTION: and Method for Use as Insecticides  
NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.

STREET: 816 Congress Avenue, Suite 1900  
CITY: Austin

STATE: TX  
COUNTRY: USA

ZIP: 78701

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:36:25 ; Search time 54.37 Seconds

(without alignments)  
298.422 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 213

Sequence: 1 MOLPLALCLICLVHTAFRV.....KRPFRANSKANQALENAY 213

Scoring table: OLIGO  
Gapop 60.0 , Gapept 60.0

Searched: 219241 seqs, 76174552 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

PIR\_68:\*\*  
1: pirl:\*\*  
2: pirl2:\*\*  
3: pirl3:\*\*  
4: pirl4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.8	428	1 TVHUER	transforming prote
2	8	3.8	429	2 JCA4965	elk1 protein - mou
3	7	3.3	130	2 D83305	hypothetical prote
4	7	3.3	148	2 C83091	hypothetical prote
5	7	3.3	197	2 G81057	hypothetical prote
6	7	3.3	197	2 A81819	hypothetical prote
7	7	3.3	305	2 S77397	yabc protein homol
8	7	3.3	312	2 D75344	conserved hypothet
9	7	3.3	321	2 B82892	conserved hypothet
10	7	3.3	328	2 S72647	hypothetical prote
11	7	3.3	398	1 S24802	polyferredoxin 6x2
12	7	3.3	408	1 D70365	filmbrial assembly
13	7	3.3	412	1 KHHUD	cathepsin D (Ec 3.
14	7	3.3	510	2 S62901	legumin 31 precurs
15	7	3.3	512	1 WMCVEM	inclusion body mat
16	7	3.3	562	2 T05758	hypothetical prote
17	7	3.3	587	2 T16617	hypothetical prote
18	7	3.3	602	2 S47880	NADH dehydrogenase
19	7	3.3	632	1 VGVNSY	surface glycoprote
20	7	3.3	681	2 E82812	outer membrane hem
21	7	3.3	781	1 TVFPDF	protein kinase Dra
22	7	3.3	917	1 S15885	hexokinase (Ec 2.7
23	7	3.3	946	2 T16297	hypothetical prote
24	7	3.3	1940	2 A59287	myosin heavy chain
25	7	3.3	6420	2 T30283	polyketide synthas
26	6	2.8	49	2 S29215	neurotoxin Tx2 - s
27	6	2.8	65	2 B25025	malx protein - K1e
28	6	2.8	72	2 T00523	hypothetical prote
29	6	2.8	83	2 S41672	tightly associated

30	6	2.8	88	2 E86639	phosphocarrier pro
31	6	2.8	89	2 D75271	hypothetical prote
32	6	2.8	97	1 GBCH	osteocalcin precur
33	6	2.8	100	2 PC1130	insulin receptor-r
34	6	2.8	102	2 G85584	unknown protein en
35	6	2.8	102	2 T51524	hypothetical prote
36	6	2.8	106	2 A82560	50S ribosomal prot
37	6	2.8	108	2 T30659	hypothetical prote
38	6	2.8	110	1 R5EC22	ribosomal protein
39	6	2.8	110	2 H85996	50S ribosomal subu
40	6	2.8	111	2 C41839	ribosomal protein
41	6	2.8	113	2 S66512	ribosomal protein
42	6	2.8	118	2 T05520	geranylgeranylated
43	6	2.8	118	2 T15239	hypothetical prote
44	6	2.8	120	2 E96808	protein F28K19.8 l
45	6	2.8	121	2 D86783	50S ribosomal prot
46	6	2.8	124	1 R3EC12	ribosomal protein
47	6	2.8	124	2 JH0443	ribosomal protein
48	6	2.8	124	2 E85998	30S ribosomal subu
49	6	2.8	124	4 JH0807	ribosomal protein
50	6	2.8	124	4 JH0809	ribosomal protein

#### ALIGNMENTS

```
RESULT 1
TVHUER
transforming protein elk-1 - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C:Accession: A41354; S54721
R: Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.
Science 244, 66-70, 1989
A:Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near transloc
A:Reference number: A41354; MUID:89203250
A:Accession: A41354
A:Molecule type: mRNA
A:Residues: 1-428 <RAO>
A:Cross-references: GB:M25269; NID:9538208; PIDN:AAA52384.1; PID:9538209
R:Gille, H.; Kortzenjann, M.; Thomae, O.; Moenaw, C.; Slaughter, C.; Cobb, M.H.; Shaw,
EMBO J. 14, 951-962, 1995
A:Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and
A:Reference number: S54721; MUID:95196758
A:Accession: S54721
A>Status: preliminary
A:Molecule type: protein
A:Residues: 318-328, 'XX', 331, 336-364, 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X'
C:Genetics:
A:Gene: GDB: ELK1
A:Cross-references: GDB:119867; OMIM:311040
A:Map position: Xp11.2-Xp11.2
A:Superfamily: elk-1-transforming protein; ets DNA-binding domain homology
C:Keywords: DNA binding; oncogene; transforming protein
F:7-86/Domain: ets DNA-binding domain homology <ETS>
```

Query Match 3.8%; Score 8; DB 1; Length 428;

Best local similarity 100.0%; Pred. No. 3.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKGRKPR 196  
Db 310 POKGRKPR 317

```
RESULT 2
JCA4965
elk1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 16-Jul-1999
C:Accession: JCA4965; I48339; I48340; S54908
R:Grevin, D.; Ung, S.; Denhez, F.; Denhem, M.; Quatanens, B.; Begue, A.; Stehelin, D.
```

Gene 174, 185-188, 1996  
A:Title: Structure and organization of the mouse elk1 gene.  
A:Reference number: J04965; MUID:97017146  
A:Accession: J04965  
A:Molecule type: mRNA  
A:Residues: 1-429 <GRE>  
A:Cross-references: EMBL:X87257; NID:9836634; PIDN:CAA60715.1; PID:9836635  
A:Experimental source: embryo  
R:Giovane, A.; Pintzas, A.; Maira, S.M.; Sobiesczuk, P.; Waslylyk, B.  
Gene Dev. 8, 1502-1513, 1994  
A:Title: Net, a new ets transcription factor that is activated by Ras.  
A:Reference number: A53837; MUID:95047310  
A:Accession: I48339  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-429 <RES>  
A:Cross-references: EMBL:X87257; NID:9836634; PIDN:CAA60715.1; PID:9836635  
A:Accession: I48340  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 5-132, 'T', 134-224 <RE2>  
A:Cross-references: EMBL:Z36939; NID:9535922; PIDN:CAA85391.1; PID:9535923  
A:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which has a DNA-binding domain homology <ETS>  
C:Genetics:  
A:Gene: elk1  
A:Introns: 70/3; 219/3; 363/3; 397/3  
C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology  
F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 3.8%; Score 8; DB 2; Length 429;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 POKGRPR 196  
|||||||  
DB 311 POKGRPR 318

RESULT 3  
D83305  
hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83305  
R:Stover, C.K.; Pham, X.Q.; Errin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Badian, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.  
A:Reference number: A82950; MUID:20437337  
A:Accession: D83305  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-130 <STO>  
A:Cross-references: GB:AE004700; GB:AE004091; NID:9948792; PIDN:AA606110.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2722

Query Match 3.3%; Score 7; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KVRIVAS 164  
|||||||  
DB 12 KVRIVAS 18

RESULT 4  
C83091  
hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: C83091  
R:Stover, C.K.; Pham, X.Q.; Errin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Badian, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.  
A:Reference number: A82950; MUID:20437337  
A:Accession: C83091  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-148 <STO>  
A:Cross-references: GB:AE004858; GB:AE004091; NID:9950668; PIDN:AA607829.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4441

Query Match 3.3%; Score 7; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 112 ARLLPNA 118  
|||||||  
DB 26 ARLLPNA 32

RESULT 5  
G81057  
hypothetical protein NMB1656 [imported] - Neisseria meningitidis (strain MC58 serogroup B)  
C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: G81057  
R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755  
A:Accession: G81057  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <DET>  
A:Cross-references: GB:AE002516; GB:AE002098; NID:97226905; PIDN:AA642005.1; PID:97222  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1656

Query Match 3.3%; Score 7; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 203 KANQAEI 209  
|||||||  
DB 122 KANQAEI 128

RESULT 6  
A81819  
hypothetical protein NMA1913 [imported] - Neisseria meningitidis (strain 22491 serogroup B)  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: A81819  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; MO; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491  
A:Reference number: A8175; MUID:2022556  
A:Accession: A81819  
A:Status: preliminary

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:54 ; Search time 31.03 Seconds  
(without alignments) 235.141 Million cell updates/sec

Title: US-09-668-021-6  
Perfect score: 213  
Sequence: 1 MDPLALCLICLIVHFAFRV.....KRRPRASAKAQAELNAY 213

Scoring table:  
Gapop 60.0 , Gapext 60.0

Word size : 0  
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.8	428	1 ELK1_HUMAN	P19419 homo sapien
2	8	3.8	429	1 ELK1_MOUSE	P11669 mus muscula
3	7	3.3	110	1 RL22_LEPIN	Q9X031 leptocystira
4	7	3.3	305	1 YB44_STRY3	P73460 synechocyst
5	7	3.3	412	1 CARD_HUMAN	P07339 homo sapien
6	7	3.3	512	1 IBMP_FMYD	P09524 figwort mos
7	7	3.3	602	1 NUSM_DIDNA	P41309 didelphis m
8	7	3.3	632	1 VGLG_SYNY	P27277 sonchus yel
9	7	3.3	781	1 KRAF_DROME	P11346 drosophila
10	7	3.3	917	1 HXK2_MOUSE	O08528 mus muscula
11	7	3.3	917	1 HXK2_RAT	P27881 rattus norv
12	6	2.8	49	1 TX25_PHONI	P29424 phoneutria
13	6	2.8	65	1 MALX_KLEPN	P29852 klebsiella
14	6	2.8	97	1 OSTC_CHICK	P02822 gallus gall
15	6	2.8	97	1 VET_HPV44	O080914 human papil
16	6	2.8	110	1 RL22_ECOLI	P02423 escherichia
17	6	2.8	111	1 RL22_ACHLA	P29222 acholeplasma
18	6	2.8	112	1 RL22_SPICT	O31160 spiroplasma
19	6	2.8	113	1 RL22_THERH	P48266 thermus aqu
20	6	2.8	115	1 TINF_HUMAN	O05411 homo sapien
21	6	2.8	115	1 TINF_MOUSE	Q92154 mus musculu
22	6	2.8	118	1 REV_HVILW	Q70624 human immun
23	6	2.8	123	1 RS12_ECOLI	P02367 escherichia
24	6	2.8	131	1 RL22_PHYS1	O66094 phytoplasm
25	6	2.8	167	1 B3AR_MERON	O70432 meriones un
26	6	2.8	168	1 TRIC_CHICK	P27673 gallus gall
27	6	2.8	175	1 RUGA_YEAST	O02326 saccharomyc
28	6	2.8	182	1 YCRO_YEAST	P21653 saccharomyc
29	6	2.8	196	1 RETB_CHICK	P41263 gallus gall
30	6	2.8	206	1 BTF3_HUMAN	P20290 homo sapien
31	6	2.8	210	1 TRPF_KLOUA	P13997 kluyveromyc
32	6	2.8	211	1 YH73_METTH	O71801 methanobact
33	6	2.8	213	1 PYRE_HAFTN	P43855 haemophilus

## ALIGNMENTS

RESULT	ID	ELK1_HUMAN	STANDARD:	PRT:	428 AA.	
34	6	2.8	216	1	EGFH_HUMAN	O60258 homo sapien
35	6	2.8	216	1	EGFH_MOUSE	O70627 mus musculu
36	6	2.8	216	1	YE97_MYCPN	P75290 mycoplasma
37	6	2.8	230	1	219_HUMAN	P98173 homo sapien
38	6	2.8	237	1	NRL_MOUSE	P53223 saccharomyc
39	6	2.8	239	1	YGIP_YEAST	O29491 macroscelid
40	6	2.8	240	1	DHAN_MACPR	O29491 macroscelid
41	6	2.8	247	1	PS72_XENLA	O99Y61 xenopus lae
42	6	2.8	248	1	PS71_XENLA	O99Y61 xenopus lae
43	6	2.8	248	1	PSA7_HUMAN	O14818 homo sapien
44	6	2.8	248	1	PSA7_MOUSE	O92200 mus musculu
45	6	2.8	249	1	APX1_PEA	P48534 pisum sativ
46	6	2.8	249	1	PSA7_CHICK	O13268 gallus gall
47	6	2.8	254	1	PSA7_RAT	P48004 rattus norv
48	6	2.8	263	1	PANB_BUCAL	P57293 buchnera ap
49	6	2.8	263	1	TRUA_BACHD	O92910 bacillus ha
50	6	2.8	268	1	ISPE_AQUAE	O67060 aquilex aeo

RESULT 1  
ELK1\_HUMAN  
ID P19419; 075606; 09UIMA; 095058; PRT; 428 AA.  
AC 01-NOV-1990 (Rel. 16, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE ETS-DOMAIN PROTEIN ELK-1.  
GN ELK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=89203250; PubMed=2539641;  
RA Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,  
RA Reddy E.S.P.;  
RT "elk, tissue-specific ets-related genes on chromosomes X and 14 near  
translocation breakpoints.";  
RL Science 244:66-70(1989).  
RN [2]  
RP MEDLINE=99013876; PubMed=9795224;  
RA Harindranath N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;  
RT "The human elk-1 gene family: the functional gene and two processed  
pseudogenes embedded in the Igh locus.";  
RL Gene 221:215-224(1998).  
RN [3]  
RP MEDLINE=92215224(1998).  
RA Graffham D.;  
RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP MEDLINE=92334979; PubMed=1630903;  
RA Janknecht R., Nordheim A.;  
RT "Elk-1 protein domains required for direct and SRF-assisted  
DNA-binding.";  
RL Nucleic Acids Res. 20:3317-3324(1992).  
RN [5]  
RP MEDLINE=92334979; PubMed=1630903;  
RA Janknecht R., Nordheim A.;  
RT "Elk-1 protein domains required for direct and SRF-assisted  
DNA-binding.";  
RL Nucleic Acids Res. 20:3317-3324(1992).  
CC -!- FUNCTION: STIMULATES TRANSCRIPTION, BINDS TO PURINE-RICH DNA  
SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE  
ELEMENT.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE), AND 2/ELKV; ARE

```

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: LUNG AND TESTIS.
CC -1- PRT: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (338, JNK,
CC AND ERK).
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M25269; AAA52384.1; -
CC EMBL; AF080616; AAC82466.1; -
CC EMBL; AL009172; CAA15659.1; -
CC EMBL; AF000672; AAD00862.1; -
CC PIR; A41354; TVHUER.
CC HSSP; P14921; 2STW.
CC TRANSFAC; T00250; -
CC MIM; 311040; -
CC InterPro; IPR000418; -
CC Pfam; PF00178; Ets; 1.
CC PRINTS; PR00454; ETSDOMAIN.
CC PROSITE; PS00345; ETS_DOMAIN_1; 1.
CC PROSITE; PS00346; ETS_DOMAIN_2; 1.
CC PROSITE; PS50061; ETS_DOMAIN_3; 1.
CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC Phosphorylation; Alternative splicing.
CC DNA_BIND 5 86 ETS-DOMAIN.
CC FT VARSPPLIC 91 95 VAGCS -> SHCAP (IN ISOFORM 2).
CC FT CONFLICT 96 428 MISSING (IN ISOFORM 2).
CC FT CONFLICT 183 183 S -> N (IN REF. 1).
CC SEQUENCE 428 AA; 4488 MW; 68F71F8ADB9D38CA CRC64;

Query Match 3.88; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 POKGRKPR 196
DB 310 POKGRKPR 317

RESULT 2
ELK1_MOUSE STANDARD; PRT; 429 AA.
P41969;
01-NOV-1995 (Rel. 32, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
ETS-DOMAIN PROTEIN ELK-1.
DE ELK1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RA Grevin D., Ung S., Denhez F., Dehem M., Quatennens B., Begue A.,
RA Stehelin D., Martin P.;
RT "Structure and organization of the mouse elk1 gene.";
RL Gene 174:185-188(1996).
RN [2]
RP SEQUENCE OF 5-224 FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=95047310; PubMed=7958835;
RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Wasyluk B.;
RT "Net, a new ets transcription factor that is activated by Ras.";
```

```

RL Genes Dev. 8:1502-1513(1994).
CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRP MOTIFS OF THE FOS SERUM RESPONSE
CC ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X87257; CAA60715.1; -
CC EMBL; Z36939; CAA85391.1; -
CC HSSP; P14921; 2STW.
CC MGD; MG1:101833; ELK1.
CC InterPro; IPR000418; -
CC Pfam; PF00178; Ets; 1.
CC PRINTS; PR00454; ETSDOMAIN.
CC PROSITE; PS00345; ETS_DOMAIN_1; 1.
CC PROSITE; PS00346; ETS_DOMAIN_2; 1.
CC PROSITE; PS50061; ETS_DOMAIN_3; 1.
CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC Phosphorylation.
CC DNA_BIND 5 86 ETS-DOMAIN.
CC FT CONFLICT 133 133 P -> F (IN REF. 2).
CC SEQUENCE 429 AA; 45243 MW; B61B5B977731D54F CRC64;

Query Match 3.88; Score 8; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 POKGRKPR 196
DB 311 POKGRKPR 318

RESULT 3
RL22_LEPIN STANDARD; PRT; 110 AA.
AC Q9XD31;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L22.
GN RPLV.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROVAR LAI;
RX MEDLINE=20088835; PubMed=10620683;
RA Zuercher R.L., Hartskeel R.A., van de Kemp H., Bal A.E.;
RT "Characterization of the Leptospira interrogans S10-spc-alpha
RT operon.";
RL FEBS Microbiol. Lett. 182:303-308(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S RNA; ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:17 ; Search time 89.8 seconds

(without alignments)  
313.819 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 213  
Sequence: 1 MQLPLALCLICLLVHTAFRV.....KPPRRARSAKANQALEENAY 213

Scoring table:  
Gapop 60.0 , Gapext 60.0

Word size : 0  
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

1: SP-archaea:\*  
2: SP-bacteria:\*  
3: SP-fungi:\*  
4: SP-human:\*  
5: SP-invertebrate:\*  
6: SP-mammal:\*  
7: SP-mhc:\*  
8: SP-organelle:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-rodent:\*  
12: SP-unclassified:\*  
13: SP-vertebrate:\*  
14: SP-virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	198	10	Q9FYQ3
2	8	3.8	276	13	Q9YGD6
3	8	3.8	276	13	Q9PT38
4	8	3.8	276	13	Q9PT37
5	8	3.8	276	13	Q9PT36
6	8	3.8	1064	4	Q75770
7	8	3.8	1064	4	Q9NQ38
8	7	3.3	89	10	Q9FPH5
9	7	3.3	109	5	Q9GTU4
10	7	3.3	130	2	Q910C0
11	7	3.3	148	2	Q9HVK4
12	7	3.3	186	10	Q9SDR7
13	7	3.3	197	2	Q9JYB9
14	7	3.3	197	2	Q9JYB0
15	7	3.3	208	2	Q9JYB1
16	7	3.3	251	2	Q9JYB2
17	7	3.3	282	10	Q9XFI9
18	7	3.3	312	2	Q9RT99
19	7	3.3	321	2	Q9P072

20	7	3.3	398	1	Q00388	Q00388 methanococ
21	7	3.3	408	2	Q66951	Q66951 aquifex aeo
22	7	3.3	425	5	Q19348	Q19348 caenorhabd
23	7	3.3	457	14	Q9WRM6	Q9WRM6 macaca mla
24	7	3.3	510	10	Q39521	Q39521 cryptomeria
25	7	3.3	546	13	Q91A33	Q91A33 pinephales
26	7	3.3	562	10	Q81837	Q81837 arabidopsis
27	7	3.3	587	5	Q21432	Q21432 caenorhabd
28	7	3.3	604	5	Q44003	Q44003 toxoplasma
29	7	3.3	638	4	Q9HY77	Q9HY77 homo sapien
30	7	3.3	648	4	Q9HAI2	Q9HAI2 homo sapien
31	7	3.3	667	4	Q9Y2N7	Q9Y2N7 homo sapien
32	7	3.3	681	2	Q9PCB8	Q9PCB8 xylella fas
33	7	3.3	739	5	Q9W4Z3	Q9W4Z3 drosophila
34	7	3.3	782	5	Q9NEH9	Q9NEH9 drosophila
35	7	3.3	802	5	Q96398	Q96398 schistosoma
36	7	3.3	815	4	Q43273	Q43273 homo sapien
37	7	3.3	917	11	Q54892	Q54892 rattus norv
38	7	3.3	946	5	Q20143	Q20143 caenorhabd
39	7	3.3	980	4	Q9NS55	Q9NS55 homo sapien
40	7	3.3	1045	4	Q9UNR9	Q9UNR9 homo sapien
41	7	3.3	1045	4	Q9NS56	Q9NS56 homo sapien
42	7	3.3	1325	2	Q9W6X3	Q9W6X3 acetobacter
43	7	3.3	1935	5	Q44934	Q44934 loligo peal
44	7	3.3	1940	5	Q02456	Q02456 schistosoma
45	7	3.3	2517	5	Q9W060	Q9W060 drosophila
46	7	3.3	6420	2	P95814	P95814 streptomyce
47	6	2.8	22	2	Q85607	Q85607 streptomyce
48	6	2.8	38	4	Q9UN17	Q9UN17 homo sapien
49	6	2.8	40	14	Q92266	Q92266 simlan herp
50	6	2.8	62	10	Q9LDG4	Q9LDG4 oryza sativ

## ALIGNMENTS

RESULT 1  
ID Q9FYQ3 PRELIMINARY; PRT; 198 AA.  
AC Q9FYQ3;  
DT 01-MAR-2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae;  
OC Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
clone: P0433F09."  
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF002539; BAB08182.1; -  
SQ SEQUENCE 198 AA; 21843 MW; 11DDID90FF4B72F4E CRC64;

Query Match 3.8%; Score 8; DB 10; Length 198;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 PGGEAPRA 156  
DB 83 PGGEAPRA 90  
RESULT 2  
Q9YGD6 PRELIMINARY; PRT; 276 AA.  
AC Q9YGD6;

DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.  
 DR EMBL: AF100930; AAD20217.1; -.  
 DR HSSP: P50163; 2AEL.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDHRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR OXIDOREDUCTASE.  
 SQ SEQUENCE 276 AA; 30111 MW; 135604A38D2940CD CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATE 35  
 Db 93 AFKNDATE 100

RESULT 3  
 Q9PT38 PRELIMINARY; PRT; 276 AA.  
 ID Q9PT38  
 AC Q9PT38;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.  
 DR EMBL: AF100933; AAD20992.1; -.  
 DR HSSP: P50163; 2AEL.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDHRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR OXIDOREDUCTASE.  
 SQ SEQUENCE 276 AA; 30239 MW; 134B52054C751C28 CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATE 35  
 Db 93 AFKNDATE 100

RESULT 4  
 Q9PT37 PRELIMINARY; PRT; 276 AA.  
 ID Q9PT37  
 AC Q9PT37;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.  
 DR EMBL: AF100932; AAD20991.1; -.  
 DR HSSP: P50163; 2AEL.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDHRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR OXIDOREDUCTASE.  
 SQ SEQUENCE 276 AA; 30127 MW; 1350E4211D34225D CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATE 35  
 Db 93 AFKNDATE 100

RESULT 5  
 Q9PT36 PRELIMINARY; PRT; 276 AA.  
 ID Q9PT36  
 AC Q9PT36;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:34:28 ; Search time 79.75 Seconds  
(without alignments)  
161.917 Million cell updates/sec

Title: US-09-668-021-6  
Perfect score: 213  
Sequence: 1 MQLPLALCLICLVHTAFRV.....KPPRRARSAKANQAEIENAY 213

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

word size :

Total number of hits satisfying chosen parameters: 412676

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : A\_Geneseq\_0601: \*

1	/SID58/gcgdata/geneseq/geneseqp/AA1980.DAT *
2	/SID58/gcgdata/geneseq/geneseqp/AA1981.DAT *
3	/SID58/gcgdata/geneseq/geneseqp/AA1982.DAT *
4	/SID58/gcgdata/geneseq/geneseqp/AA1984.DAT *
5	/SID58/gcgdata/geneseq/geneseqp/AA1985.DAT *
6	/SID58/gcgdata/geneseq/geneseqp/AA1986.DAT *
7	/SID58/gcgdata/geneseq/geneseqp/AA1987.DAT *
8	/SID58/gcgdata/geneseq/geneseqp/AA1988.DAT *
9	/SID58/gcgdata/geneseq/geneseqp/AA1989.DAT *
10	/SID58/gcgdata/geneseq/geneseqp/AA1990.DAT *
11	/SID58/gcgdata/geneseq/geneseqp/AA1991.DAT *
12	/SID58/gcgdata/geneseq/geneseqp/AA1992.DAT *
13	/SID58/gcgdata/geneseq/geneseqp/AA1993.DAT *
14	/SID58/gcgdata/geneseq/geneseqp/AA1994.DAT *
15	/SID58/gcgdata/geneseq/geneseqp/AA1995.DAT *
16	/SID58/gcgdata/geneseq/geneseqp/AA1996.DAT *
17	/SID58/gcgdata/geneseq/geneseqp/AA1997.DAT *
18	/SID58/gcgdata/geneseq/geneseqp/AA1998.DAT *
19	/SID58/gcgdata/geneseq/geneseqp/AA1998.DAT *
20	/SID58/gcgdata/geneseq/geneseqp/AA1999.DAT *
21	/SID58/gcgdata/geneseq/geneseqp/AA2000.DAT *
22	/SID58/gcgdata/geneseq/geneseqp/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	213	100.0	213	21	AAV96430	Human TGF-beta bin
2	203	95.3	213	21	AAV96436	Human TGF-beta bin
3	175	82.2	213	21	AAAB26106	Human DAN/Cerberus
4	175	82.2	213	21	AAV96429	Human TGF-beta bin
5	175	82.2	213	22	AAV97589	Human secreted pro
6	139	65.3	367	21	AAAB26105	Human DAN/Cerberus
7	113	53.1	213	21	AAV96431	Verret TGF-beta bi
8	76	35.7	176	21	AAV96434	Bovine TGF-beta bi
9	33	15.5	213	21	AAV96433	Rat TGF-beta dindi
10	32	15.0	211	21	AAV96432	Murine TGF-beta bi
11	27	12.7	50	20	AAV12009	Human 5' EST secre

12	13	6.1	23	21	AAV66435	Mutant human TGF
13	13	4.2	116	18	AAW27654	Secreted protein A
14	9	4.2	116	18	AAW44090	Human secreted pro
15	9	4.2	206	19	AAW09408	Human small CCN-II
16	9	4.2	206	19	AAW58704	Human small CCN-II
17	9	4.2	206	21	AAW65711	Homo sapiens fetal
18	9	4.2	206	21	AAW10233	Human adult retina
19	9	4.2	206	21	AAW75981	Murine skin cell p
20	9	4.2	206	21	AAW76031	Murine skin cell p
21	9	4.2	206	22	AAW55920	SKin cell protein,
22	9	4.2	206	22	AAW55970	skin cell protein,
23	8	3.8	177	20	AAW33001	Human serine prote
24	8	3.8	428	20	AAW32874	Human ELK-1 protei
25	8	3.8	428	20	AAW52704	Human ELK-1, Homo
26	8	3.8	554	20	AAW6226	Human secreted pro
27	8	3.8	922	20	AAW33002	Human serine prote
28	8	3.8	1064	22	AAW6568	Serine protease in
29	8	3.8	1064	22	AAW6568	Serine protease in
30	8	3.3	61	21	AAW00348	Human secreted pro
31	7	3.3	173	21	AAW43980	Human cancer associ
32	7	3.3	186	19	AAW65018	Porphyria divergent
33	7	3.3	219	21	AAW4285	Arabidopsis thalia
34	7	3.3	265	21	AAW4284	Arabidopsis thalia
35	7	3.3	282	21	AAW4283	Arabidopsis thalia
36	7	3.3	285	21	AAW9625	Cat flea HMT perit
37	7	3.3	412	16	AAW4207	Human death associ
38	7	3.3	412	19	AAW1369	Death associated p
39	7	3.3	412	20	AAW06478	Human tumour assoc
40	7	3.3	412	21	AAW3685	Amino acid sequenc
41	7	3.3	426	22	AAW4892	Human secreted pro
42	7	3.3	457	21	AAW53197	Macaca mulatta rha
43	7	3.3	718	19	AAW72072	HSV-2 strain SB5 C
44	7	3.3	917	18	AAW37437	Rat hexokinase II.
45	7	3.3	917	18	AAW37429	Rat hexokinase II.
46	6	2.8	917	18	AAW3793	AS-30D tumour Type
47	6	2.8	10	12	AAW10268	Human ventricular
48	6	2.8	15	19	AAW45818	Peptide recognisese
49	6	2.8	15	19	AAW45613	Peptide recognisese
50	6	2.8	19	19	AAW12476	Human neuroendociti

## ALIGNMENTS

RESULT	1
AAV966430	

AA  
AC  
AA96430;

DE Human TGF-beta binding protein (BEER) variant V10I

KW osteopaplic; transforming growth factor-beta; TGF-beta; binding protein;  
 KW BEER; variant; V10I; gene therapy; antisense therapy; fracture;  
 KW chromosome 17q12-21; bone mineralization.

Key	Location/Qualifiers
PH	Misc-difference 10
FT	/label= v101
FT	/note= "wild type valine has been substituted with isoleucine"
FT	

MO200032773-A1.

08-JUN-2000.

24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.  
PA  
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
PI Van Ness J, Winkler DG;  
XX  
DR MPI: 2000-412321/35.  
DR N-PSDB; AAA29056.  
XX  
PT Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
PS Claim 3; Page 119-120; 16zpp; English.

This shows a variant human transforming growth factor-beta (TGF-beta) binding protein designated BEER V101, which comprises a substitution of isoleucine for the wild-type valine at residue 10. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

Sequence 213 AA:  
XQ  
XX

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:35:22 ; Search time 47.66 Seconds  
(without alignments)  
100.571 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 213  
Sequence: 1 MQLPALCLICLVHTAFRV.....KPRPRASAKANQALEENAV 213

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

arched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/5A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/5B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	4.2	206	1 US-08-468-847B-2	Sequence 2, Appl1
2	9	4.2	206	1 US-08-468-847B-20	Sequence 20, Appl1
3	9	4.2	206	4 US-09-188-930-159	Sequence 159, App
4	9	4.2	206	4 US-09-188-930-286	Sequence 286, App
5	7	3.3	186	4 US-09-475-316A-13	Sequence 13, Appl
6	7	3.3	412	1 US-08-208-007A-12	Sequence 12, Appl
7	7	3.3	412	4 US-08-974-691-4	Sequence 4, Appl1
8	7	3.3	917	2 US-08-588-983-16	Sequence 16, Appl
9	7	3.3	917	2 US-08-588-976-16	Sequence 16, Appl
10	6	2.8	39	3 US-08-554-385-22	Sequence 22, Appl
11	6	2.8	41	4 US-08-905-223-400	Sequence 400, App
12	6	2.8	68	2 US-08-836-791-5	Sequence 5, Appl1
13	6	2.8	77	1 US-08-370-225-13	Sequence 13, Appl
14	6	2.8	77	1 US-08-461-859-13	Sequence 13, Appl
15	6	2.8	77	5 PCT-US83-10069-13	Sequence 13, Appl
16	6	2.8	80	1 US-08-011-396B-15	Sequence 15, Appl
17	6	2.8	80	1 US-08-464-051-15	Sequence 15, Appl
18	6	2.8	80	2 US-08-462-498-15	Sequence 15, Appl
19	6	2.8	80	3 US-08-554-385-14	Sequence 14, Appl
20	6	2.8	88	2 US-08-690-011A-31	Sequence 31, Appl
21	6	2.8	97	2 US-08-245-511-30	Sequence 30, Appl
22	6	2.8	97	2 US-08-600-993A-30	Sequence 30, Appl
23	6	2.8	186	1 US-07-860-981-5	Sequence 5, Appl1
24	6	2.8	186	5 PCT-US83-09634-5	Sequence 5, Appl1
25	6	2.8	216	2 US-08-821-637-3	Sequence 3, Appl1
26	6	2.8	223	4 US-08-928-941D-22	Sequence 22, Appl
27	6	2.8	223	4 US-09-280-590A-22	Sequence 22, Appl

28	6	2.8	243	2 US-08-319-376-2	Sequence 2, Appl1
29	6	2.8	248	2 US-08-701-935-3	Sequence 3, Appl1
30	6	2.8	248	3 US-09-134-591-3	Sequence 3, Appl1
31	6	2.8	254	3 US-08-701-935-6	Sequence 6, Appl1
32	6	2.8	254	3 US-09-134-591-6	Sequence 6, Appl1
33	6	2.8	269	2 US-08-727-311-3	Sequence 3, Appl1
34	6	2.8	330	4 US-09-188-930-144	Sequence 144, App
35	6	2.8	330	4 US-09-188-930-278	Sequence 278, App
36	6	2.8	343	2 US-08-933-750C-13	Sequence 13, Appl
37	6	2.8	343	4 US-09-214-613-13	Sequence 13, Appl
38	6	2.8	354	2 US-08-700-013B-9	Sequence 9, Appl1
39	6	2.8	364	1 US-08-680-726A-56	Sequence 56, Appl1
40	6	2.8	364	4 US-09-092-409-56	Sequence 56, Appl1
41	6	2.8	372	4 US-08-928-941D-16	Sequence 16, Appl1
42	6	2.8	372	4 US-09-280-590A-16	Sequence 16, Appl1
43	6	2.8	386	3 US-08-972-902-3	Sequence 3, Appl1
44	6	2.8	388	1 US-08-087-772A-2	Sequence 2, Appl1
45	6	2.8	392	1 US-08-271-354-11	Sequence 11, Appl1
46	6	2.8	392	2 US-08-565-861-11	Sequence 11, Appl1
47	6	2.8	392	5 PCT-US94-07638-11	Sequence 11, Appl1
48	6	2.8	395	3 PCT-US94-07638-11	Sequence 11, Appl1
49	6	2.8	395	4 US-08-981-825-6	Sequence 6, Appl1
50	6	2.8	400	1 US-07-916-901-6	Sequence 6, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-468-847B-2  
; Sequence 2, Application US/08468847B  
; Patent No. 5780263  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
; TITLE OF INVENTION: Human Ccn-Like Growth Factor  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: CARELITA, BYRNE, BAIN, GILFILLIAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,847B  
; FILING DATE: 6 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MILLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-442  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-08-468-847B-2

```
Query Match          4.2%: Score 9; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36
   |||||||
Db 23 AFKNDATET 31

RESULT 2
US-08-468-847B-20
; Sequence 20, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468, 847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-468-847B-20

Query Match          4.2%: Score 9; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36
   |||||||
Db 23 AFKNDATET 31

RESULT 3
US-09-188-930-159
; Sequence 159, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James Greg

; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188, 930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: mouse
; US-09-188-930-159

Query Match          4.2%: Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36
   |||||||
Db 23 AFKNDATET 31

RESULT 4
US-09-188-930-286
; Sequence 286, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188, 930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 286
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
; US-09-188-930-286

Query Match          4.2%: Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36
   |||||||
Db 23 AFKNDATET 31

RESULT 5
US-09-475-316A-13
; Sequence 13, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Salkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
; FILE REFERENCE: MSUR-1-13793
; CURRENT APPLICATION NUMBER: US/09/475, 316A
; CURRENT FILING DATE: 1999-12-30
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:36:26 ; Search time 54.37 Seconds

(without alignments)  
298.422 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 213

Sequence: 1 MQLPLALCLVCLVIAHAFRV.....KPRPRAGAKANQALEMAY 213

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Word size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database: PIR\_68:\*

1: PIR1:\*\n2: PIR2:\*\n3: PIR3:\*\n4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3.8	428	1	TYHUPK	transforming prote
2	3.8	429	2	JC4965	elk1 protein - mou
3	3.8	850	2	JC5700	Erbb kinase activa
4	3.3	99	1	MGGPB2	beta-2-microglobul
5	3.3	119	2	S14292	transcription acti
6	3.3	130	2	D83305	hypothetical prote
7	3.3	148	2	C83091	hypothetical prote
8	3.3	183	2	B37410	H-2 class II hist
9	3.3	197	2	C81057	hypothetical prote
10	3.3	197	2	A81819	hypothetical prote
11	3.3	230	2	I37095	gene 2.19 protein
12	3.3	321	2	B82892	conserved hypotet
13	3.3	328	2	S72647	hypothetical prote
14	3.3	398	1	S24802	polyferredoxin 6x2
15	3.3	408	2	D70365	limbal assembly
16	3.3	412	1	KHHUD	cathepsin D (EC 3
17	3.3	459	2	S13064	ID-myo-inositol-tr
18	3.3	461	2	JN0129	ID-myo-inositol-tr
19	3.3	510	2	S62901	legumin 31 precurs
20	3.3	562	2	T05758	hypothetical prote
21	3.3	573	2	B70942	hypothetical prote
22	3.3	587	2	T16617	hypothetical prote
23	3.3	602	2	S47880	NADH dehydrogenase
24	3.3	632	1	VGVSX	surface glycoprote
25	3.3	681	2	E82812	outer membrane hem
26	3.3	829	2	E64114	translation initia
27	3.3	864	2	JS0076	regulatory protein
28	3.3	900	2	B70694	probable Infb - My
29	3.3	917	1	S15885	hexokinase (EC 2.7

#### ALIGNMENTS

30	7	3.3	962	2	A81817	translation initia
31	7	3.3	962	2	C81060	translation initia
32	7	3.3	1168	1	MMAXIC	myosin heavy chain
33	7	3.3	1308	2	T15280	hypothetical prote
34	7	3.3	1940	2	A59287	myosin heavy chain
35	7	3.3	6420	2	T30283	polyketide synthas
36	6	2.8	22	2	JP0066	ribosomal protein
37	6	2.8	49	2	S29215	neurotoxin Tx2 - s
38	6	2.8	65	2	B25025	malX protein - Kle
39	6	2.8	83	2	S41672	lightly associated
40	6	2.8	89	2	D75271	hypothetical prote
41	6	2.8	101	2	I57492	apolipoprotein C-I
42	6	2.8	102	2	G85584	unknown protein en
43	6	2.8	102	2	T51524	hypothetical prote
44	6	2.8	106	2	A82560	50S ribosomal prot
45	6	2.8	108	2	T49144	hypothetical prote
46	6	2.8	110	1	R5EC22	ribosomal protein
47	6	2.8	110	2	H85996	50S ribosomal subu
48	6	2.8	111	2	C41839	ribosomal protein
49	6	2.8	113	2	S6512	ribosomal protein
50	6	2.8	118	2	T05520	geranylgeranylated

RESULT 1  
TYHUPK  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1993 #sequence-revision 31-Mar-1993 #text-change 18-Jun-1999  
C:Accession: A41354; S54721  
R: Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.  
Science 244, 66-70, 1989  
A:Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near transloc  
A:Reference number: A41354; MUID:89203250  
A:Accession: A41354  
A:Molecule type: mRNA  
A:Residues: 1-428 <RMO>  
A:Cross-references: GB:M25269; NID:g538208; PIDN:AA52384.1; PID:g538209  
R:Gille, H.; Korteljan, M.; Thoma, O.; Moomaw, C.; Slaught, C.; Cobb, M.H.; Shaw,  
EMBO J. 14, 951-962, 1995  
A:Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and  
A:Reference number: S54721; MUID:95196758  
A:Accession: S54721  
A:Molecule type: protein  
A:Status: preliminary  
A:Residues: 318-328, 'XX', 331, 336-364; 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X'  
A:Gene: GDB:ELK1  
A:Cross-references: GDB:119867; OMIM:311040  
C:Map position: Xp11.2-Xp11.2  
C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology  
C:Keywords: DNA binding; oncogene; transforming protein  
F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 3.8%; Score 8; DB 1; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKGRKR 196  
| | | | | | | | | |  
DB 310 POKGRKR 317

RESULT 2  
JC4965  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1996 #sequence-revision 31-Dec-1996 #text-change 16-Jul-1999  
C:Accession: JC4965; I48339; I48340; S54908  
R:Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatennens, B.; Begue, A.; Stehelin, D.

Gene 174, 185-188, 1996  
 A:Title: Structure and organization of the mouse elk1 gene.  
 A:Reference number: Jc4965; MUID:97017146  
 A:Accession: Jc4965  
 A:Molecule type: mRNA  
 A:Residues: 1-429 <GRE>  
 A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635  
 A:Experimental source: embryo  
 R:Giovane, A.; Placzas, A.; Maiba, S.M.; Sobieszczyk, P.; Wasyluk, B.  
 Genes Dev. 8, 1502-1513, 1994  
 A:Title: Net, a new ets transcription factor that is activated by Ras.  
 A:Reference number: A53837; MUID:95047310  
 A:Accession: I48339  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-429 <RES>  
 A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635  
 A:Accession: I48340  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 5-132, 134-224 <RE2>  
 A:Cross-references: EMBL:Z36939; NID:g535922; PIDN:CAA85391.1; PID:g535923  
 A:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which is C:Genetics:  
 A:Gene: elk1  
 A:Introns: 70/3; 219/3; 363/3; 397/3  
 C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology  
 F:/-86/Domain: ets DNA-binding domain homology <ETS>  
 Query Match 3.8%; Score 8; DB 2; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 189 PQGRKPR 196  
 DB 311 PQGRKPR 318  
 RESULT 3  
 Jc5700  
 ERB kinase activator alpha, brain and thymus - human  
 C:Species: Homo sapiens (man)  
 C:Date: 25-Nov-1997 #sequence\_revision 25-Nov-1997 #text\_change 21-Jul-2000  
 C:Accession: Jc5700  
 R:Higashiyama, S.; Horikawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.; Miyeg  
 J. Biochem. 122, 675-680, 1997  
 A:Title: A novel brain-derived member of the epidermal growth factor family that interact  
 A:Reference number: Jc5700; MUID:98006324  
 A:Accession: Jc5700  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-850 <HIG>  
 A:Cross-references: DDBJ:AB005060; NID:g2626738; PIDN:BAA23417.1; PID:g2626739  
 A:Experimental source: SK-NSH cell  
 C:Comment: This protein is a member of the epidermal growth factor family. It is functio  
 ating the differentiation of MDA-MB-453 cells.  
 C:Superfamily: unassigned EGF-related proteins; EGF homology  
 C:Keywords: glycoprotein  
 F:258-311/Domain: 19-116 #status predicted <IGL>  
 F:345-381/Domain: EGF homology <EGF>  
 F:346-381/Domain: EGF-like #status predicted <EGF2>  
 F:147,278,451/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 3.8%; Score 8; DB 2; Length 850;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 195 PRPRARGA 202  
 DB 725 PRPRARGA 732

RESULT 4  
 MCGPB2  
 beta-2-microglobulin - guinea pig  
 N:Alternate names: class I histocompatibility antigen beta chain  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 31-Jan-1997  
 C:Accession: A02181  
 R:Molte, P.B.; Gebra, J.J.  
 Mol. Immunol. 17, 1493-1505, 1980  
 A:Title: The primary structure of guinea pig beta2-microglobulin.  
 A:Reference number: A02181; MUID:82057805  
 A:Accession: A02181  
 A:Molecule type: protein  
 A:Residues: 1-99 <MOL>  
 C:Complex: heterodimer with MHC class I histocompatibility antigen alpha chain  
 C:Function:  
 A:Description: necessary for the expression of MHC class I histocompatibility antigen  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterodimer  
 F:18-82/Domain: immunoglobulin homology <IMM>  
 F:25-80/Disulfide bonds: #status predicted  
 Query Match 3.3%; Score 7; DB 1; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 LTVHNAF 18  
 DB 64 LTVHNAF 70  
 RESULT 5  
 S14292  
 transcription activator - Emericella nidulans  
 C:Species: Emericella nidulans, Aspergillus nidulans  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 02-Jul-1998  
 C:Accession: S14292  
 R:Kulmburg, P.; Prange, T.; Mathieu, M.; Sequeval, D.; Sczacocchio, C.; Felenbok, B.  
 FEBS Lett. 280, 11-16, 1991  
 A:Title: Correct intron splicing generates a new type of a putative zinc-binding dom  
 A:Reference number: S14292; MUID:91184391  
 A:Accession: S14292  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-119 <KUL>  
 C:Superfamily: GAL4 zinc binuclear cluster homology  
 F:/-54/Domain: GAL4 zinc binuclear cluster homology <GALA>  
 Query Match 3.3%; Score 7; DB 2; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 151 GAAPRAR 157  
 DB 63 GAAPRAR 69  
 RESULT 6  
 DB8305  
 hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: DB8305  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: AB2950; MUID:20437337  
 A:Accession: DB8305

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:55 ; Search time 31.03 Seconds  
(without alignments)  
235.141 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 213  
Sequence: 1 MQLPLALCLVLCILVHAFRV.....KRPFRARAKANQAELENNY 213

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Indexed: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.8	428	1 ELK1_HUMAN	P19419 homo sapien
2	8	3.8	429	1 ELK1_MOUSE	P41969 mus musculu
3	8	3.8	850	1 NRG2_HUMAN	O14511 homo sapien
4	7	3.3	99	1 B2MG_CAVPO	P01886 cavia porce
5	7	3.3	110	1 RL22_LEPIN	Q9X431 leptospira
6	7	3.3	230	1 219_HUMAN	P98173 homo sapien
7	7	3.3	412	1 CATD_HUMAN	P07339 homo sapien
8	7	3.3	459	1 IP3K_RAT	P17105 rattus norv
9	7	3.3	461	1 IP3K_HUMAN	P23677 homo sapien
10	7	3.3	602	1 N03M_DIDMA	P41309 didelphis m
11	7	3.3	632	1 VGLG_SYNY	P27277 sonchus yel
12	7	3.3	821	1 ALCR_EMENT	P21228 emeritella
13	7	3.3	829	1 IP2_HAEIN	P43233 haemophilus
14	7	3.3	900	1 IP2_MYCTU	P71613 mycobacteri
15	7	3.3	917	1 HXK2_MOUSE	O08528 mus musculu
16	7	3.3	917	1 HXK2_RAT	P27881 rattus norv
17	7	3.3	924	1 IP2_MTCLE	Q9Z519 mycobacteri
18	7	3.3	1168	1 MTSQ_ACACA	P01559 acanthamoeb
19	6	2.8	49	1 TX25_PHONI	P29424 phoneutria
20	6	2.8	53	1 LHA1_RHOAC	P35089 rhodopsendo
21	6	2.8	65	1 MALX_KLEPN	P28892 klebsiella
22	6	2.8	97	1 VER2_HPV4	O080914 human papil
23	6	2.8	101	1 APC2_MACFA	P18658 macaca fasc
24	6	2.8	110	1 RL22_ECOLI	P24243 escherichia
25	6	2.8	111	1 RL22_ACHLA	P29222 acholeplasm
26	6	2.8	112	1 RL22_SPICI	O31160 spiroplasma
27	6	2.8	113	1 RL22_THERH	P48286 thermus aqu
28	6	2.8	115	1 TIAF_HUMAN	O95414 homo sapien
29	6	2.8	115	1 TIAF_MOUSE	O92154 mus musculu
30	6	2.8	118	1 REV_HVILM	Q70624 human immun
31	6	2.8	119	1 Y13K_NPVOP	O05122 oryza pseu
32	6	2.8	123	1 RS12_ECOLI	P02367 escherichia
33	6	2.8	131	1 RL22_PHYS1	O66094 phytoplasma

34	6	2.8	135	1 I14_BOVIN	P30367 bos taurus
35	6	2.8	135	1 I14_CAPIT	P79155 capra hircu
36	6	2.8	135	1 I14_SHEEP	P30368 ovis aries
37	6	2.8	146	1 Y1IR_ECOLI	P32161 escherichia
38	6	2.8	148	1 RK22_MAIZE	P06589 zea mays (m
39	6	2.8	149	1 RK22_ORYZA	P12140 oryza sativ
40	6	2.8	150	1 R19E_PYRHO	O59041 pyrococcus
41	6	2.8	167	1 B3AR_MERUN	O70432 meriones un
42	6	2.8	168	1 TRIC_CHICK	P27673 gallus gall
43	6	2.8	175	1 RU6A_YEAST	Q02326 saccharomyc
44	6	2.8	182	1 YCYO_YEAST	P25654 saccharomyc
45	6	2.8	193	1 HS72_CANAL	P46587 candida alb
46	6	2.8	196	1 RENB_CHICK	P41263 gallus gall
47	6	2.8	210	1 TRPF_KLUTA	P13997 kiuyveromyc
48	6	2.8	213	1 PYRE_HAEIN	P43855 haemophilus
49	6	2.8	216	1 FGPH_HUMAN	O60258 homo sapien
50	6	2.8	216	1 FGPH_MOUSE	O70627 mus musculu

## ALIGNMENTS

RESULT	ID	ELK1_HUMAN	STANDARD:	PRT:	428 AA.
AC	1	ELK1_HUMAN	P19419; O75606; O9UIM4; O95058;		
DT	01-NOV-1990	(Rel. 16, Created)			
DT	01-OCT-2000	(Rel. 40, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	ETS-DOMAIN PROTEIN ELK-1.				
GN	ELK1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=89203250; PubMed=2539641;				
RX	Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,				
RA	Reddy E.S.P.;				
RT	"elk, tissue-specific ets-related genes on chromosomes X and 14 near				
RT	translocation breakpoints.";				
RL	Science 244:66-70(1989).				
RN	[2]	SEQUENCE FROM N.A.			
RP	MEDLINE=99013876; PubMed=9795224;				
RX	Hardstrandath N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;				
RA	"The human elk-1 gene family: the functional gene and two processed				
RT	pseudogenes embedded in the 19h locus.";				
RL	Gene 221:215-224(1998).				
RN	[3]	SEQUENCE FROM N.A.			
RP	Grafham D.;				
RA	Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.				
RL	[4]	SEQUENCE FROM N.A. (ISOFORM 2).			
RP	Arjee D.N.T., Kovar H.;				
RA	"Novel family members HUEP1, ELFR, and ELKY among ETS-related genes				
RT	coexpressed with EMS-FL11 in Ewing tumor cell lines.";				
RL	Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.				
RN	[5]	DOMAINS.			
RP	MEDLINE=92334979; PubMed=1630903;				
RA	Jankecht R., Nordheim A.;				
RT	"Elk-1 protein domains required for direct and SRF-assisted				
RL	DNA-binding.";				
CC	Nucleic Acids Res. 20:3317-3324(1992).				
CC	-I- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA				
CC	SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE				
CC	FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE				
CC	ELEMENT.				
CC	-I- SUBCELLULAR LOCATION: NUCLEAR.				
CC	-I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE), AND 2/ELKY; ARE				

```

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: LONG AND TESTIS.
CC -1- PWM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,
CC AND ERK).
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M25269; AAC52384.1; -.
CC EMBL: AF080616; AAC82466.1; -.
CC EMBL: AL009172; CAAL5659.1; -.
CC EMBL: AF000672; AAD00862.1; -.
CC PIR: A41354; TVHUK.
CC HSSP: P14921; 2STW.
CC TRANSFAC: T00250; -.
CC MIM: 311040; -.
CC InterPro: IPR000418; -.
CC Pfam: PF00178; Ets; 1.
CC PRINTS: PR00454; ETSDOMAIN.
CC PROSITE: PS00345; ETS_DOMAIN_1; 1.
CC PROSITE: PS00346; ETS_DOMAIN_2; 1.
CC PROSITE: PS50061; ETS_DOMAIN_3; 1.
CC DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC KW Phosphorylation; Alternative splicing.
CC FT DNA_BIND 5 86 ETS-DOMAIN.
CC FT VARSPPLIC 91 95 VAGCS -> SHCAP (IN ISOFORM 2).
CC FT VARSPPLIC 96 428 MISSING (IN ISOFORM 2).
CC FT CONFLICT 183 183 S -> N (IN REF. 1).
CC FT SEQUENCE 428 AA; 44888 MW; 68F71F8ADB9D38CA CRC64;
SQ

```

Query Match 3.8%; Score 8; DB 1; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 189 POKGRKPR 196
DB 310 POKGRKPR 317

```

```

RESULT 2
ELK1_MOUSE
ID ELK1_MOUSE STANDARD; PRT; 429 AA.
AC P41969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ETS-DOMAIN PROTEIN ELK-1.
GN ELK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97017146; PubMed=8863747;
RA Grevin D., Ung S., Denhez F., Dehem M., Quatanens B., Begue A.,
RA Stehelin D., Martin P.;
RT "Structure and organization of the mouse elk1 gene.";
RL Gene 174:185-188(1996).
RN [2]
RP SEQUENCE OF 5-224 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95047310; PubMed=7958835;
RA Giovane A., Pintzas A., Maira S.-M., Sobiesczuk P., Wasyluk B.;
RT "Net, a new ets transcription factor that is activated by Ras.";

```

```

RL Genes Dev. 8:1502-1513(1994).
CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
CC ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X87257; CAA60715.1; -.
CC EMBL: Z36939; CAA85391.1; -.
CC HSSP: P14921; 2STW.
CC MGD: MGI:101833; ELK1.
CC InterPro: IPR000418; -.
CC Pfam: PF00178; Ets; 1.
CC PRINTS: PR00454; ETSDOMAIN.
CC PROSITE: PS00345; ETS_DOMAIN_1; 1.
CC PROSITE: PS00346; ETS_DOMAIN_2; 1.
CC PROSITE: PS50061; ETS_DOMAIN_3; 1.
CC DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC KW Phosphorylation.
CC FT DNA_BIND 5 86 ETS-DOMAIN.
CC FT CONFLICT 133 133 P -> T (IN REF. 2).
CC FT SEQUENCE 429 AA; 45243 MW; B61B5B97731D54F CRC64;
SQ

```

Query Match 3.8%; Score 8; DB 1; Length 429;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 189 POKGRKPR 196
DB 311 POKGRKPR 318

```

```

RESULT 3
NRG2_HUMAN
ID NRG2_HUMAN STANDARD; PRT; 850 AA.
AC O14511;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRO-NEUREGULIN-2 PRECURSOR (PRO-NRG2) [CONTAINS: NEUREGULIN-2 (NRG-2)
DE (NEURAL-AND THYMUS-DERIVED ACTIVATOR FOR ERBB KINASES) (NTAK)
DE (DIVERGENT OF NEUREGULIN 1) (DON-1)].
GN NRG2 OR NTAK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Neuroblastoma;
RX MEDLINE=98006324; PubMed=9348101;
RA Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,
RA Nakagawa T., Miyagawa J., Matsushita N., Negatsu T., Taniguchi N.,
RA Ishiguro H.;
RT "A novel brain-derived member of the epidermal growth factor family
RT that interacts with ErbB3 and ErbB4.";
RL J. Biochem. 122:675-680(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS DON-1B AND DON-1R).
RC TISSUE=Fetal brain;
RX MEDLINE=97342638; PubMed=9199335;

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:18 ; Search time 89.8 Seconds  
(without alignments)  
313.819 Million cell updates/sec

Title: US-09-668-021-10  
Perfect score: 213  
Sequence: 1 M0PLALCVLLVHAAFRV.....KPRPARAKANQAELENNV 213

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Archived: 425026 seqs, 133305027 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

SPREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	276	13	Q9YGD6 oncorhynchu
2	8	3.8	276	13	Q9PT38 oncorhynchu
3	8	3.8	276	13	Q9PT37 oncorhynchu
4	8	3.8	276	13	Q9PT36 oncorhynchu
5	7	3.3	45	4	Q9HC30 homo sapien
6	7	3.3	89	10	Q9FH55 aradidopsis
7	7	3.3	109	5	Q9GT34 ancylostoma
8	7	3.3	114	5	Q9GT34 ancylostoma
9	7	3.3	125	11	Q9EQ01 Q9eq1 cavia porce
10	7	3.3	130	2	Q9I0C0 pseudomonas
11	7	3.3	148	2	Q9HVX4 pseudomonas
12	7	3.3	150	2	Q9EUV6 pasteurella
13	7	3.3	151	2	Q9EUV4 actinobacill
14	7	3.3	151	2	Q9EUV2 haemophilus
15	7	3.3	151	2	Q9EUV1 haemophilus
16	7	3.3	151	2	Q9EUV0 haemophilus
17	7	3.3	151	2	Q9EUV9 haemophilus
18	7	3.3	151	2	Q9EUV8 haemophilus
19	7	3.3	151	2	Q9EUV6 haemophilus

20	7	3.3	151	2	Q9EUV5 haemophilus
21	7	3.3	151	2	Q9EUV4 haemophilus
22	7	3.3	151	2	Q9EUV3 haemophilus
23	7	3.3	151	2	Q9EUV2 haemophilus
24	7	3.3	151	2	Q9EUV1 haemophilus
25	7	3.3	151	2	Q9EUV0 haemophilus
26	7	3.3	151	2	Q9EUV9 haemophilus
27	7	3.3	151	2	Q9EUV8 haemophilus
28	7	3.3	151	2	Q9EUV7 haemophilus
29	7	3.3	151	2	Q9EUV6 haemophilus
30	7	3.3	151	2	Q9EUV5 haemophilus
31	7	3.3	151	2	Q9EUV4 haemophilus
32	7	3.3	151	2	Q9EUV3 haemophilus
33	7	3.3	151	2	Q9EUV2 haemophilus
34	7	3.3	151	2	Q9EUV1 haemophilus
35	7	3.3	151	2	Q9EUV0 haemophilus
36	7	3.3	151	2	Q9EUV9 haemophilus
37	7	3.3	151	2	Q9EUV8 haemophilus
38	7	3.3	151	2	Q9EUV7 haemophilus
39	7	3.3	151	2	Q9EUV6 haemophilus
40	7	3.3	151	2	Q9EUV5 haemophilus
41	7	3.3	151	2	Q9EUV4 haemophilus
42	7	3.3	151	2	Q9EUV3 haemophilus
43	7	3.3	151	2	Q9EUV2 haemophilus
44	7	3.3	151	2	Q9EUV1 haemophilus
45	7	3.3	151	2	Q9EUV0 haemophilus
46	7	3.3	151	2	Q9EUV9 haemophilus
47	7	3.3	151	2	Q9EUV8 haemophilus
48	7	3.3	151	2	Q9EUV7 haemophilus
49	7	3.3	151	2	Q9EUV6 haemophilus
50	7	3.3	151	2	Q9EUV5 haemophilus

#### ALIGNMENTS

RESULT 1  
Q9YGD6 PRELIMINARY; PRT: 276 AA.  
AC Q9YGD6; 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Procanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-99185307; PubMed=10082666;  
RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
RT "Cloning and expression of two carbonyl reductase-like 20beta-  
hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
trout (Oncorhynchus mykiss).";  
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
FAMILY.  
CC EMBL; AF100930; AAD20217.1; -.  
DR HSSP; P50163; ZAE1.  
DR InterPro; IPR002347; -.  
DR pfam; PF00106; adh\_short; 1.  
DR PRINTS; PRO0081; GDHRDH.  
DR PRINTS; PRO0080; SDRFAMILY.  
KW Oxidoreductase.  
SQ SEQUENCE 276 AA; 30111 MW; 135604A38D2940CD CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATE 35  
 |||||||  
 DB 93 AFKNDATE 100

RESULT 2  
 O9PT38 PRELIMINARY; PRT; 276 AA.

AC 09PT38  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBOXYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostomi;  
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NX NCBI\_TaxID=8022;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 RL trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.

DR EMBL: AF100933; AAD20992.1; -.  
 DR HSSP: P50163; 2AE1.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDRHDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SO SEQUENCE 276 AA; 30239 MW; 134B52054C751C28 CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATE 35  
 |||||||  
 DB 93 AFKNDATE 100

RESULT 3  
 O9PT37 PRELIMINARY; PRT; 276 AA.

AC 09PT37  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBOXYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostomi;  
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NX NCBI\_TaxID=8022;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 RL trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.  
 DR EMBL: AF100932; AAD20991.1; -.

DR HSSP: P50163; 2AE1.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDRHDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SO SEQUENCE 276 AA; 30127 MW; 1350E4211D34225D CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATE 35  
 |||||||  
 DB 93 AFKNDATE 100

RESULT 4  
 O9PT36 PRELIMINARY; PRT; 276 AA.

AC 09PT36  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CARBOXYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostomi;  
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NX NCBI\_TaxID=8022;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 RL trout (Oncorhynchus mykiss).";  
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.

DR EMBL: AF100931; AAD20218.1; -.  
 DR HSSP: P50163; 2AE1.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDRHDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SO SEQUENCE 276 AA; 30139 MW; 135F1711D30086D CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATE 35  
 |||||||  
 DB 93 AFKNDATE 100

RESULT 5  
 O9HC30 PRELIMINARY; PRT; 45 AA.

AC 09HC30  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CAMP SPECIFIC PHOSPHODIESTERASE VARIANT PDE4A-10 (FRAGMENT).  
 GN PDE4A.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:34:29 ; Search time 79.75 Seconds  
(without alignments)  
161.917 Million cell updates/sec

Title: US-09-668-021-10  
Perfect score: 213  
Sequence: 1 MQLPLALCIVCLVHNAFRV.....KPPRRARAGAKAQALENNAY 213

Scoring table:  
Gapop 60.0 , Gapext 60.0

Word size : 0  
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

A\_Geneseq.0601.\*  
1: /SID8/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID8/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID8/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID8/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID8/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID8/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID8/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID8/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID8/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID8/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID8/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID8/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID8/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID8/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID8/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID8/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID8/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID8/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID8/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID8/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID8/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID8/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	100.0	213	21	AAV96431
2	135	63.4	213	21	AAV96431
3	135	63.4	213	21	AAV96431
4	135	63.4	213	21	AAV96431
5	115	54.0	176	21	AAV96431
6	113	53.1	213	21	AAV96431
7	113	53.1	213	21	AAV96431
8	77	36.2	367	21	AAV96431
9	33	15.5	50	20	AAV12009
10	33	15.5	213	21	AAV96431
11	32	15.0	211	21	AAV96431

12	15	7.0	23	21	AAV96435
13	9	4.2	116	18	AAV27654
14	9	4.2	116	18	AAV44090
15	9	4.2	206	18	AAV09408
16	9	4.2	206	19	AAV58704
17	9	4.2	206	20	AAV95711
18	9	4.2	206	21	AAV10233
19	9	4.2	206	21	AAV75981
20	9	4.2	206	21	AAV76031
21	9	4.2	206	22	AAV58920
22	9	4.2	206	22	AAV58970
23	8	3.8	428	20	AAV52704
24	8	3.8	428	20	AAV32824
25	8	3.8	647	19	AAV48383
26	8	3.3	55	21	AAV40943
27	7	3.3	61	21	AAV00348
28	7	3.3	79	20	AAV73811
29	7	3.3	99	21	AAV68229
30	7	3.3	99	21	AAV52883
31	7	3.3	99	22	AAV58644
32	7	3.3	114	19	AAV68587
33	7	3.3	114	19	AAV75969
34	7	3.3	173	21	AAV43980
35	7	3.3	177	20	AAV33801
36	7	3.3	186	19	AAV65018
37	7	3.3	219	21	AAV42285
38	7	3.3	230	21	AAV57845
39	7	3.3	256	20	AAV88386
40	7	3.3	265	21	AAV42284
41	7	3.3	282	21	AAV42283
42	7	3.3	285	21	AAV29625
43	7	3.3	412	16	AAV74207
44	7	3.3	412	19	AAV71369
45	7	3.3	412	20	AAV06478
46	7	3.3	412	21	AAV93685
47	7	3.3	426	22	AAV64892
48	7	3.3	554	20	AAV36226
49	7	3.3	708	20	AAV88383
50	7	3.3	917	18	AAV37437

#### ALIGNMENTS

RESULT 1	
AAV96431	standard; Protein: 213 AA.
AC	AAV96431:
DT	12-SEP-2000. (first entry)
DE	Vervet TGF-beta binding protein (BEER).
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; gene therapy; antisense therapy; fracture; bone mineralization.
OS	Cercopithecus pygerythrus.
PN	WC200032773-A1.
PD	08-JUN-2000.
PE	24-NOV-1999: 99WO-US27990.
PR	27-NOV-1998: 98US-0110283.
PA	(DARW-) DARWIN DISCOVERY LTD.
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
PL	Van Ness J, Winkler DG;
DR	WPI: 2000-412321/35.

Mutant human TGF- $\beta$   
Secreted protein A  
Human secreted pro  
Human small CN-11  
Homo sapiens fetal  
Human adult retina  
Murine skin cell p  
Skin cell protein,  
Skin cell protein,  
Human ELK-1 protel  
Homo sapiens don-1  
Human ORFX ORF707  
Human secreted pro  
Human prostate tum  
Guinea pig beta-2-  
Guinea pig beta-2-  
Guinea pig beta 2m  
Tobacco hawkmoth E  
Ecdisis trilingering  
Human cancer assoc  
Human setine prote  
Forsythia dirigent  
Arabidopsis thalia  
Human cytokine fam  
Mammalian Zneul po  
Arabidopsis thalia  
Arabidopsis thalia  
Cat flea HMT perit  
Human death associ  
Death associated p  
Human tumour-assoc  
Amino acid sequenc  
Human secreted pro  
Mammalian Zneul po  
Rat hexokinase II.

DR N-PSDB; AAA29057.  
 XX  
 PT Nucleic acids (I) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 XX  
 PS Claim 4; Page 122-123; 162pp; English.  
 XX  
 CC This shows a vervet transforming growth factor-beta (TGF-beta)  
 CC binding protein designated vBEER. The CDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BEER expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta Bp expression. The CDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BEER by  
 CC expressing inactive proteins or to supplement the patient's own production  
 CC of BEER polypeptides. The nucleic acids may be used for recombinant  
 CC production of BEER, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BEER may be used to raise  
 CC antibodies and for identification of BEER modulators. BEER antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 CC  
 XX  
 SQ Sequence 213 AA:

Query Match 100.0%; Score 213; DB 21; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 4, 3e-214;  
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLPLAICVCLLVHAAFRVGGGQMAFKNDATETIPELGYPEPPELENKTKMRAE 60  
 DB 1 MQLPLAICVCLLVHAAFRVGGGQMAFKNDATETIPELGYPEPPELENKTKMRAE 60  
 QY 61 NGRPRPHHPEETKDVSSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCCPARILPNAIG 120  
 DB 61 nggrprphhpeetkdvsseyscrelhfttryvtdgpcrsakpytelvcsqgcparrilpnaig 120  
 QY 121 RGMKWRSPGPPFCIPDRYRAQRVQLLCPGGAAPRAKRVLVASCCKRLTRFHNSGLK 180  
 DB 121 rgmkwrspgppfcipdryraqrvqlcpggaapraakrvlvascckrltrfhngselk 180  
 QY 181 DFGPEAARPOKRRKPRPARGAKANOAELENAY 213  
 DB 181 dfgpeaarpgkrrkprpargakangaelenay 213

## RESULT 2

AAB26106  
 ID AAB26106 standard; Protein: 213 AA.  
 XX

AC AAB26106;  
 DT 15-JAN-2001 (first entry)  
 DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).  
 KW Human: DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;  
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;  
 KW gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055193-A2.  
 PD 21-SEP-2000.  
 PF 02-MAR-2000; 2000WO-US05537.  
 PR 12-MAR-1999; 99US-0124118.  
 PA (REGG-) REGENERON PHARM INC.

XX Economides AN.  
 PI  
 DR WPI: 2000-638179/61.  
 DR N-PSDB; AAA94051.  
 XX

PT Novel isolated, human DNA/Cerberus related protein 6 which include  
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic  
 PT acids encoding the proteins which are useful as probes and primers  
 XX  
 PS Claim 8; Fig 3; 40pp; English.  
 XX

CC The present sequence comprises the amino acid sequence encoded by exons 1  
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding  
 CC sequence. The coding sequence was isolated from a human kidney cDNA  
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely  
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of  
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and  
 CC protein can be used as immunogens, modulators of cell function, growth  
 CC and differentiation, to reduce undesirable bone formation, to identify  
 CC DCR6 binding agents, in diagnosis, and in gene therapy.  
 CC  
 XX

SQ Sequence 213 AA:

Query Match 63.4%; Score 135; DB 21; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 9, 7e-133;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AFRVVGQGMQAFKNDATETIPELGYPEPPELENKTKMRAENGRRPHHPEYDVS 76  
 DB 17 afrvvgqgwmqafkndateilpeilgypeppelelnktmraengrrphhpeftkds 76  
 QY 77 EYSCRELHFTRYVTDGPCRSAPVTELVCSGCCPARILPNAIGRGMKWRSPGPPFCIP 136  
 DB 77 eyscrelhfttryvtdgpcrsakpytelvcsqgcparrilpnaigrkwrspgppfcip 136  
 QY 137 DRYRAQRVQLLCPGG 151  
 DB 137 dryraqrvqlcpgg 151

## RESULT 3

AAAY96429  
 ID AAAY96429 standard; Protein: 213 AA.  
 XX

AC AAAY96429;  
 DT 12-SEP-2000 (first entry)  
 DE Human TGF-beta binding protein (BEER).  
 DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200032773-A1.  
 PD 08-JUN-2000.  
 PF 24-NOV-1999; 99WO-US27990.  
 PR 27-NOV-1998; 98US-0110283.  
 PA (DARW-) DARWIN DISCOVERY LTD.  
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
 PI Van Ness J, Winkler DG;  
 DR WPI: 2000-412321/35.  
 DR N-PSDB; AAA29055.  
 XX

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:35:23 : Search time 47.66 Seconds  
(without alignments)  
100.571 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 213  
Sequence: 1 MOLPLALCVCLVHAAPRY.....KPRPRAGAKANOETENAY 213

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Word size: 0  
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.2	206	1	US-08-468-847B-2
2	9	4.2	206	1	US-08-468-847B-20
3	9	4.2	206	4	US-09-188-930-159
4	9	4.2	206	4	US-09-188-930-286
5	8	3.8	647	3	US-08-753-007A-32
6	8	3.8	647	4	US-09-398-496-32
7	7	3.3	99	2	US-08-484-905-61
8	7	3.3	99	3	US-08-481-985B-61
9	7	3.3	99	4	US-08-370-476-61
10	7	3.3	114	1	US-08-582-257-15
11	7	3.3	114	2	US-08-582-298-15
12	7	3.3	186	4	US-09-475-316A-13
13	7	3.3	412	1	US-08-208-007A-12
14	7	3.3	412	4	US-08-974-691-4
15	7	3.3	861	1	US-08-764-100-24
16	7	3.3	917	2	US-08-588-983-16
17	7	3.3	917	2	US-08-588-976-16
18	6	2.8	39	3	US-08-554-385-22
19	6	2.8	68	2	US-08-836-791-5
20	6	2.8	80	1	US-08-011-398B-15
21	6	2.8	80	1	US-08-464-051-15
22	6	2.8	80	2	US-08-462-498-15
23	6	2.8	80	3	US-08-554-385-14
24	6	2.8	88	2	US-08-690-011A-31
25	6	2.8	97	2	US-08-245-511-30
26	6	2.8	97	2	US-08-600-993A-30
27	6	2.8	125	1	US-08-225-989-20

28	6	2.8	125	1	US-08-570-923-20	Sequence 20, Appl
29	6	2.8	125	1	US-08-580-014-20	Sequence 20, Appl
30	6	2.8	125	3	US-08-985-526-25	Sequence 25, Appl
31	6	2.8	125	4	US-09-079-785-20	Sequence 20, Appl
32	6	2.8	128	1	US-08-666-798-2	Sequence 2, Appl
33	6	2.8	128	1	US-08-892-692-2	Sequence 2, Appl
34	6	2.8	128	2	US-09-096-071-2	Sequence 2, Appl
35	6	2.8	163	4	US-09-354-129-12	Sequence 12, Appl
36	6	2.8	186	1	US-07-960-981-5	Sequence 5, Appl
37	6	2.8	186	5	PCT-US93-09634-5	Sequence 5, Appl
38	6	2.8	199	3	US-08-737-248-7	Sequence 7, Appl
39	6	2.8	202	4	US-09-342-084-10	Sequence 10, Appl
40	6	2.8	215	1	US-08-223-989-23	Sequence 23, Appl
41	6	2.8	215	1	US-08-570-923-23	Sequence 23, Appl
42	6	2.8	215	1	US-08-580-014-23	Sequence 23, Appl
43	6	2.8	215	4	US-09-079-785-23	Sequence 23, Appl
44	6	2.8	216	2	US-08-821-637-3	Sequence 3, Appl
45	6	2.8	223	4	US-08-928-941D-22	Sequence 22, Appl
46	6	2.8	223	4	US-09-280-590A-22	Sequence 22, Appl
47	6	2.8	234	1	US-08-225-989-8	Sequence 8, Appl
48	6	2.8	234	1	US-08-570-923-8	Sequence 8, Appl
49	6	2.8	234	1	US-08-580-014-8	Sequence 8, Appl
50	6	2.8	234	4	US-09-079-785-8	Sequence 8, Appl

#### ALIGNMENTS

RESULT 1  
US-08-468-847B-2  
: Sequence 2, Application US/08468847B  
: Patent No. 5780263  
: GENERAL INFORMATION:  
: APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
: TITLE OF INVENTION: Human CCN-Like Growth Factor  
: NUMBER OF SEQUENCES: 20  
: CORRESPONDENCE ADDRESS:  
: ADDRESSER: CARELIA, BYRNE, BAIN, GILFILLAN,  
: ADDRESSER: CECCHI, STEWART & OLSTEIN  
: STREET: 6 BECKER FARM ROAD  
: CITY: ROSELAND  
: STATE: NEW JERSEY  
: COUNTRY: USA  
: ZIP: 07068  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3.5 INCH DISKETTE  
: COMPUTER: IBM PS/2  
: OPERATING SYSTEM: MS-DOS  
: SOFTWARE: WORD PERFECT 5.1  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/468,847B  
: FILING DATE: 6 June 1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER:  
: FILING DATE:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: MULLINS, J.G.  
: REGISTRATION NUMBER: 33,073  
: REFERENCE/DOCKET NUMBER: 325800-442  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 201-994-1700  
: TELEFAX: 201-994-1744  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 206 AMINO ACIDS  
: TYPE: AMINO ACID  
: STRANDEDNESS: LINEAR  
: TOPOLOGY: LINEAR  
: MOLECULE TYPE: PROTEIN  
: US-08-468-847B-2

Query Match 4.2%; Score 9; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

RESULT 2  
US-08-468-847B-20  
Sequence 20, Application US/08468847B  
Patent No. 5780263  
GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ. ID NO.: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-468-847B-20

Query Match 4.2%; Score 9; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

RESULT 3  
US-09-188-930-159  
Sequence 159, Application US/09188930A  
Patent No. 6150502  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James Greg

TITLE OF INVENTION: Compositions Isolated From Skin Cells  
TITLE OF INVENTION: and Methods For Their Use  
FILE REFERENCE: 11000.1011c1  
CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ. ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ. ID NO. 159  
LENGTH: 206  
TYPE: PRT  
ORGANISM: mouse  
US-09-188-930-159

Query Match 4.2%; Score 9; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

RESULT 4  
US-09-188-930-286  
Sequence 286, Application US/09188930A  
Patent No. 6150502  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James Greg  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
TITLE OF INVENTION: and Methods For Their Use  
FILE REFERENCE: 11000.1011c1  
CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ. ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ. ID NO. 286  
LENGTH: 206  
TYPE: PRT  
ORGANISM: Mouse  
US-09-188-930-286

Query Match 4.2%; Score 9; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

RESULT 5  
US-08-753-007A-32  
Sequence 32, Application US/08753007A  
Patent No. 6074841  
GENERAL INFORMATION:  
APPLICANT: Gearing, David P.  
APPLICANT: Busfield, Samantha J.  
TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES  
TITLE OF INVENTION: AND USES THEREFOR  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:36:28 ; Search time 54.37 Seconds

(without alignments)  
295.619 Million cell updates/sec

Title: US-09-668-021-12

Perfect score: 211

Sequence: 1 MOPSLAPCLICLVHAFCA.....KPPGARGAKANQALEENAY 211

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

arched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

PIR 68:\*  
1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	428	1	TVHUEK
2	8	3.8	429	2	JCA4965
3	7	3.3	99	1	MGCPB2
4	7	3.3	130	2	D83305
5	7	3.3	148	2	C83091
6	7	3.3	197	2	G81057
7	7	3.3	197	2	A81819
8	7	3.3	222	2	B41643
9	7	3.3	222	2	B55356
10	7	3.3	321	2	B82892
11	7	3.3	327	2	A55356
12	7	3.3	328	2	S42152
13	7	3.3	328	2	S72647
14	7	3.3	354	2	C83577
15	7	3.3	358	1	B24802
16	7	3.3	531	2	B83082
17	7	3.3	535	1	A39194
18	7	3.3	535	1	D86028
19	7	3.3	541	2	T48517
20	7	3.3	542	2	T23869
21	7	3.3	562	2	T05758
22	7	3.3	565	2	B70652
23	7	3.3	632	1	VGVSX
24	7	3.3	681	2	E82812
25	7	3.3	836	2	B84417
26	7	3.3	852	2	A28459
27	7	3.3	1168	1	MMAXIC
28	7	3.3	1308	2	T15280
29	7	3.3	1940	2	A59287

30	6	2.8	22	2	JP0066	ribosomal protein
31	6	2.8	33	2	S43312	2',3'-cyclic-nucle
32	6	2.8	49	2	S29215	neurotoxin Tx2 - s
33	6	2.8	67	2	T30718	hypothetical prote
34	6	2.8	72	2	T00523	hypothetical prote
35	6	2.8	89	2	D75221	hypothetical prote
36	6	2.8	102	2	G85584	unknown protein en
37	6	2.8	103	2	G84293	hypothetical prote
38	6	2.8	108	1	S02802	thioredoxin C-2 -
39	6	2.8	110	1	R5EC22	ribosomal protein
40	6	2.8	110	2	H85996	50S ribosomal subu
41	6	2.8	111	2	C41839	ribosomal protein
42	6	2.8	113	2	S66512	ribosomal protein
43	6	2.8	121	2	D86783	50S ribosomal prot
44	6	2.8	123	2	S69660	hypothetical prote
45	6	2.8	130	2	S19133	hypothetical prote
46	6	2.8	134	2	D75534	ribosomal protein
47	6	2.8	136	2	T45360	hypothetical prote
48	6	2.8	138	2	D83833	hypothetical prote
49	6	2.8	142	2	JC4997	hypothetical 15.7k
50	6	2.8	142	2	H72600	hypothetical prote

#### ALIGNMENTS

RESULT 1  
TVHUEK  
Transforming protein elk-1 - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999  
C:Accession: A41354; S54721  
R: Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.  
Science 244, 66-70, 1989  
A:Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near transloc  
A:Reference number: A41354; MUID:89203250  
A:Accession: A41354  
A:Molecule type: mRNA  
A:Residues: 1-428 <RAO>  
A:Cross-references: GB:M25269; NID:9538208; PIDN:AA52384.1; PID:9538209  
R: Gilie, H.; Kortenjan, M.; Thomae, O.; Moenaw, C.; Slaughter, C.; Cobb, M.H.; Shaw,  
EMBO J. 14, 951-962, 1995  
A:Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and  
A:Reference number: S54721; MUID:95196758  
A:Accession: S54721  
A:Status: preliminary  
A:Molecule type: protein  
A:Gene: GDB:ELK1  
A:Cross-references: GDB:119867; OMIM:311040  
A:Map position: Xp11.2-Xp11.2  
A:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology  
C:Keywords: DNA binding; oncogene; transforming protein  
F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 3.8%; Score 8; DB 1; Length 428;

Best local similarity 100.0%; Pred. No. 3.1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 POKGRKPR 194

DB 310 POKGRKPR 317

RESULT 2

JCA4965

elk1 protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 16-Jul-1999  
C:Accession: JCA4965; I48339; I48340; S54908  
R:Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatanens, B.; Begue, A.; Stehelin, D.

Gene 174, 185-188, 1996  
 A:Title: Structure and organization of the mouse elk1 gene.  
 A:Reference number: J04965; MUID:97017146  
 A:Accession: J04965  
 A:Molecule type: mRNA  
 A:Residues: 1-429 <GRE>  
 A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635  
 A:Experimental source: embryo  
 R:Glovane, A.; Pitzas, A.; Maira, S.M.; Sodjeszczuk, P.; Wasylyk, B.  
 Genes Dev. 8, 1502-1513, 1994  
 A:Title: Net, a new ets transcription factor that is activated by Ras.  
 A:Reference number: A53837; MUID:95047310  
 A:Accession: I48339  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-429 <RES>  
 A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635  
 A:Accession: I48340  
 A:Molecule type: mRNA  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Residues: 5-132, 'T', 134-224 <RE2>  
 A:Cross-references: EMBL:Z50939; NID:g535922; PIDN:CAA85391.1; PID:g535923  
 C:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which h  
 C:Genetics:  
 A:Gene: elk1  
 A:Introns: 70/3; 219/3; 363/3; 397/3  
 C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology  
 F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 3.8%; Score 8; DB 2; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 POKGRKR 194  
 |||||  
 DB 311 POKGRKR 318

RESULT 3  
 MGFB2  
 beta-2-microglobulin - guinea pig  
 N:Alternate names: class I histocompatibility antigen beta chain  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 31-Jan-1997  
 C:Accession: A02181  
 R:Wolfe, P.B.; Cebra, J.J.  
 Mol. Immunol. 17, 1493-1505, 1980  
 A:Title: The primary structure of guinea pig beta2-microglobulin.  
 A:Reference number: A02181; MUID:82057805  
 A:Accession: A02181  
 A:Molecule type: protein  
 A:Residues: 1-99 <MOL>  
 C:Complex: heterodimer with MHC class I histocompatibility antigen alpha chain  
 C:Function:  
 A:Description: necessary for the expression of MHC class I histocompatibility antigen  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterodimer  
 F:18-82/Domain: immunoglobulin homology <IMM>  
 F:25-80/Disulfide bonds: #Status predicted

Query Match 3.3%; Score 7; DB 1; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LTVNAF 18  
 |||||  
 DB 64 LTVNAF 70

RESULT 4  
 DB3305

hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: D83305  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: D83305  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-130 <STO>  
 A:Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AA06110.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2722

Query Match 3.3%; Score 7; DB 2; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 KVRVAS 162  
 |||||  
 DB 12 KVRVAS 18

RESULT 5  
 C83091  
 hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: C83091  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: C83091  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <STO>  
 A:Cross-references: GB:AE004858; GB:AE004091; NID:g9950668; PIDN:AA07829.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4441

Query Match 3.3%; Score 7; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 ARLLPNA 116  
 |||||  
 DB 26 ARLLPNA 32

RESULT 6  
 G81057  
 hypothetical protein NMB1656 [imported] - Neisseria meningitidis (strain MC58 serogro  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: G81057  
 R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Frazer, C.M.; Moxon, E.R.; Rappaport, R.;  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:56 ; Search time 31.03 Seconds  
(without alignments)  
232.933 Million cell updates/sec

Title: US-09-668-021-12  
Perfect score: 211  
Sequence: 1 MQPSLAPCLICLVHAAFC.....KPRGARGAKNAQLENAY 211

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.8	428	1 ELK1_HUMAN	P19419 homo sapien
2	8	3.8	429	1 ELK1_MOUSE	P19419 mus musculus
3	7	3.3	97	1 VE7_HPV44	Q80914 human papill
4	7	3.3	99	1 B2MG_CAVPO	P01886 cavia porce
5	7	3.3	110	1 RL22_LEPIN	Q94331 leptospira
6	7	3.3	220	1 UPAS_RAT	P51573 rattus norv
7	7	3.3	222	1 UPAS_MOUSE	P35457 mus musculus
8	7	3.3	281	1 PAB4_ECOLI	P22997 escherichia
9	7	3.3	327	1 UPAR_MOUSE	P35456 mus musculus
10	7	3.3	328	1 UPAR_RAT	P49616 rattus norv
11	7	3.3	535	1 DPPA_ECOLI	P23847 escherichia
12	7	3.3	632	1 VGLG_SYNV	P27277 sonchus yel
13	7	3.3	852	1 CSG_HALHA	P08198 halobacteri
14	7	3.3	1168	1 MYSC_ACACA	P10569 acanthamoeb
15	6	2.8	49	1 TX25_PHONI	P28424 phonotria
16	6	2.8	53	1 LHAI_RHOC	P35089 rhodopsin
17	6	2.8	107	1 TH12_CORNE	P07887 corynebacte
18	6	2.8	110	1 RL22_ECOLI	P02423 escherichia
19	6	2.8	111	1 RL22_ACHLA	P29222 acholoplasm
20	6	2.8	112	1 RL22_SPICI	O31160 spiroplasma
21	6	2.8	113	1 RL22_THETH	P48286 thermus aqu
22	6	2.8	115	1 TIAF_HUMAN	O95411 homo sapien
23	6	2.8	115	1 TIAF_MOUSE	O92184 mus musculu
24	6	2.8	118	1 REV_HVILW	Q70624 human immun
25	6	2.8	131	1 RL22_PHYS1	O66094 phytoplasm
26	6	2.8	145	1 SNL_RAT	O35775 rattus norv
27	6	2.8	167	1 B3AR_MERUN	O70432 meriones un
28	6	2.8	173	1 FRIS_LYNST	P42577 lymnaea sta
29	6	2.8	182	1 YCYO_YEAST	P25654 saccharomyc
30	6	2.8	193	1 H572_CANAL	P46587 candida alb
31	6	2.8	196	1 RTB_CHICK	O41263 gallus gall
32	6	2.8	204	1 BAD_MOUSE	Q61337 mus musculu
33	6	2.8	210	1 TRPF_KLULA	P13997 klyverinomy

34	6	2.8	213	1 PYRE_HABIN	P43855 haemophilus
35	6	2.8	216	1 EGFH_HUMAN	O60258 homo sapien
36	6	2.8	216	1 EGFH_MOUSE	O70627 mus musculu
37	6	2.8	227	1 PRL_HUMAN	P01236 homo sapien
38	6	2.8	227	1 PRL_MOUSE	P55151 macaca mula
39	6	2.8	230	1 COX2_BRAFL	O47428 branchiosto
40	6	2.8	239	1 COX2_BRALA	O79417 mycobacteri
41	6	2.8	245	1 DAPB_MYCTU	P72024 mycobacteri
42	6	2.8	249	1 APX1_PEA	P48534 pisum sativ
43	6	2.8	261	1 DHSB_RICPR	Q9zeal rickettsia
44	6	2.8	263	1 LEF5_NPVOP	O10344 orgyia pseu
45	6	2.8	267	1 DAPB_BACSU	P42976 bacillus su
46	6	2.8	269	1 ESLL_MYCPN	P75333 mycoplasma
47	6	2.8	272	1 THIM_BACSU	P39593 bacillus su
48	6	2.8	280	1 HES1_HUMAN	Q14469 homo sapien
49	6	2.8	281	1 HES1_RAT	Q04666 rattus norv
50	6	2.8	282	1 HES1_MOUSE	P35428 mus musculu

#### ALIGNMENTS

```

RESULT 1
ELK1_HUMAN
ID ELK1_HUMAN STANDARD; PRT; 428 AA.
AC PI9419; O75606; Q9UJM4; O95058;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ETS-DOMAIN PROTEIN ELK-1.
GN ELK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89203250; PubMed=2539641;
RA Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,
RT Reddy E.S.P.;
RA "elk, tissue-specific ets-related genes on chromosomes X and 14 near
RT translocation breakpoints.";
RN Science 244:66-70(1989).
RP SEQUENCE FROM N.A.
RX MEDLINE=99013876; PubMed=9795224;
RA Harindranath N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;
RT "The human elk-1 gene family: the functional gene and two processed
RT pseudogenes embedded in the IgH locus.";
RN Gene 221:215-224(1998).
RP SEQUENCE FROM N.A.
RA Grafham D.;
RN Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Aryee D.N.T., Kovar H.;
RT "Novel family members HuEr71, ELFR, and ELKv among ETS-related genes
RT coexpressed with EMS-Flil in Ewing tumor cell lines.";
RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RP DOMAINS
RX MEDLINE=92334979; PubMed=1630903;
RA Janknecht R., Nordheim A.;
RT "Elk-1 protein domains required for direct and SRF-assisted
RT DNA-binding.";
RN Nucleic Acids Res. 20:3317-3324(1992).
CC -!- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
CC ELEMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE), AND 2/ELKV; ARE

```

```

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: LUNG AND TESTIS.
CC -!- PTM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,
CC AND ERK).
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25269; AAA52384.1; -.
DR EMBL; AF080616; AAC82466.1; -.
DR EMBL; AF009172; CAAL5659.1; -.
DR EMBL; AF000672; AAD00862.1; -.
DR PIR; A41354; TVHUER.
DR HSP; P14921; 2STW.
DR TRANSFAC; T00250; -.
DR MIM; 311040; -.
DR InterPro; IPR000418; -.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation; Alternative splicing.
FT DNA_BIND 5 86 ETS-DOMAIN.
FT VARSPPLIC 91 95 VAGCS -> SHCAP (IN ISOFORM 2).
FT VARSPPLIC 96 428 MISSING (IN ISOFORM 2).
FT CONFLICT 183 183 S -> N (IN REF. 1).
SQ SEQUENCE 428 AA; 44888 MW; 68F71F6ADB9D38CA CRC64;
Query Match 3.8%; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 POKGRKRP 194
DB 310 POKGRKRP 317
|||||
RESULT 2
ELK1_MOUSE
ID ELK1_MOUSE STANDARD; PRT; 429 AA.
AC F41969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ETS-DOMAIN PROTEIN ELK-1.
GN ELK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL; TISSUE=Embryo;
RX MEDLINE=97017146; PubMed=8863747;
RA Grevin D., Ung S., Denhez F., Dehem M., Quatannens B., Begue A.,
RA Stehelin D., Martin P.;
RT "Structure and organization of the mouse elk1 gene.";
RL Gene 174:185-188(1996).
RN [2]
RP SEQUENCE OF 5-224 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95047310; PubMed=7958835;
RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Wasyluk B.;
RT "Net, a new ets transcription factor that is activated by Ras.";

```

```

RL Genes Dev. 8:1502-1513(1994).
CC -!- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
CC ELEMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87257; CAA60715.1; -.
DR EMBL; Z36939; CAA85391.1; -.
DR HSP; P14921; 2STW.
DR MGD; MGI:101833; Elk1.
DR InterPro; IPR000418; -.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation.
FT DNA_BIND 5 86 ETS-DOMAIN.
FT CONFLICT 133 133 P -> T (IN REF. 2).
SQ SEQUENCE 429 AA; 45243 MW; B61B5B977731D54F CRC64;
Query Match 3.8%; Score 8; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 POKGRKRP 194
DB 311 POKGRKRP 318
|||||
RESULT 3
VE7_HPV44
ID VE7_HPV44 STANDARD; PRT; 97 AA.
AC Q80914;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 44.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
OX NCBI_TaxID=10592;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31788; AAA79458.1; -.
DR InterPro; IPR000148; -.

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:18 ; Search time 89.8 seconds  
(without alignments)  
310.873 Million cell updates/sec

Title: US-09-668-021-12  
Perfect score: 211  
Sequence: 1 MPSLAPCLICLIVHAFCA.....KPPRGARSAKANAQALENNAY 211

Scoring table:  
Gapop 60.0 , Gapext 60.0

Word size: 0  
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database: SPTREMBL.16.\*

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.3	89	09FH55	09fh55 arabidopsis
2	7	3.3	109	09GR14	09gr14 arabidopsis
3	7	3.3	125	09EQ01	09eq01 cavia porce
4	7	3.3	130	09I0C0	09i0c0 pseudomonas
5	7	3.3	148	09HVX4	09hvx4 pseudomonas
6	7	3.3	197	09JVB9	09jvb9 neisseria m
7	7	3.3	197	09JTA0	09jta0 neisseria m
8	7	3.3	208	053581	053581 streptomyces
9	7	3.3	239	09IBJ1	09ibj1 turkey herp
10	7	3.3	251	09ILNO	09ilno streptomyces
11	7	3.3	276	09KZ52	09kz52 streptomyces
12	7	3.3	282	09XF19	09xf19 arabidopsis
13	7	3.3	321	09PQ72	09pq72 ureaplasma
14	7	3.3	322	09RI19	09ri19 cricetus
15	7	3.3	328	09R119	09r119 ratulus norv
16	7	3.3	345	09FVU1	09fvu1 zea mays (m
17	7	3.3	354	09ISV7	09isv7 pseudomonas
18	7	3.3	362	09RUX3	09rux3 streptomyces
19	7	3.3	365	09LG89	09lg89 oryza sativ

20	7	3.3	398	1	000388	000388 methanococcus
21	7	3.3	425	5	019348	019348 caenorhabditis
22	7	3.3	465	10	09LTV6	09lvt6 arabidopsis
23	7	3.3	483	10	09LTK34	09ltk34 arabidopsis
24	7	3.3	486	10	09FKS0	09fks0 arabidopsis
25	7	3.3	531	2	09HVR9	09hvr9 pseudomonas
26	7	3.3	532	4	09GLK1	09glk1 homo sapien
27	7	3.3	541	10	09LYD4	09lyd4 arabidopsis
28	7	3.3	542	5	009594	009594 caenorhabditis
29	7	3.3	562	10	081837	081837 arabidopsis
30	7	3.3	565	2	P95077	p95077 mycobacterium
31	7	3.3	604	5	044003	044003 toxoplasma
32	7	3.3	638	4	09H7Y7	09h7y7 homo sapien
33	7	3.3	643	14	085050	085050 pseudorabies
34	7	3.3	681	2	09PG88	09pg88 xyloella fas
35	7	3.3	802	5	096398	096398 schistosoma
36	7	3.3	836	1	09HM69	09hm69 halobacterium
37	7	3.3	1186	5	061080	061080 acanthamoeba
38	7	3.3	1188	4	09H3X8	09h3x8 homo sapien
39	7	3.3	1308	5	001924	001924 caenorhabditis
40	7	3.3	1325	2	09WX63	09wx63 acetobacter
41	7	3.3	1935	5	044934	044934 loligo peal
42	7	3.3	1940	5	002456	002456 schistosoma
43	6	2.8	22	2	085607	085607 streptomyces
44	6	2.8	64	1	09HJ81	09hj81 thermoplasma
45	6	2.8	67	14	098283	098283 molluscum c
46	6	2.8	72	10	064466	064466 arabidopsis
47	6	2.8	73	10	022594	022594 parhatis n
48	6	2.8	81	2	09ZEL9	09zel9 lactobacillus
49	6	2.8	89	2	09RRN0	09rrn0 deinococcus
50	6	2.8	91	14	Q72747	Q72747 human immun

## ALIGNMENTS

RESULT 1  
ID 09FH55 PRELIMINARY: PRT: 89 AA.  
AC 09FH55;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K1120.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxId=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE=20181125; PubMed=10718197;  
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
RT clones." ; 7:31-63(2000).  
RL DNA Res. 7:31-63(2000).  
DR EMBL; AB022211; BAB10713.1;  
SO SEQUENCE 89 AA; 10345 MW; DF12344C836FAE0A CRC64;

Query Match 3.3%; Score 7; DB 10; Length 89;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 RLPLNAI 117  
| | | | | | | | | |  
Db 18 RLPLNAI 24

RESULT 2

Q9GTJ4 PRELIMINARY; PRT; 109 AA.  
ID Q9GTJ4  
AC Q9GTJ4:  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
OS ZINC METALLOPEPTIDASE 3 ME3 (FRAGMENT).  
OC Ancylostoma caninum (Dog hookworm).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
OX NCBI\_TaxID=29170;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=WHOLE WORKS;  
RA Jones B.F., Hotez P.J.;  
RT "Cloning and characterization of a zinc-metalloprotease secreted by  
the invasive stages of Ancylostoma caninum."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF273706; AAG29106.1; -  
FT NON\_TER  
SQ SEQUENCE 109 AA; 12300 MW; 7C3C54FD724CAD7B CRC64;

Query Match 3.3%; Score 7; DB 5; Length 109;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 129 DFRICPD 135  
Db 86 DFRICPD 92

RESULT 3  
Q9EQ01 PRELIMINARY; PRT; 125 AA.  
ID Q9EQ01  
AC Q9EQ01:  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Leclaire K.P., Naylor J.W., Kawasaki E.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF148875; AAG43561.1; -  
SQ SEQUENCE 125 AA; 13977 MW; 3F3224B02F0B41C3 CRC64;

Query Match 3.3%; Score 7; DB 11; Length 125;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LTVHAAF 18  
Db 90 LTVHAAF 96

RESULT 4  
Q910C0 PRELIMINARY; PRT; 130 AA.  
ID Q910C0  
AC Q910C0:  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA2722.  
GN PA2722.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.

OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Brody R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Gardner L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL: AE004700; AAG06110.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 130 AA; 14641 MW; ECEBDCD7F247836E CRC64;

Query Match 3.3%; Score 7; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 156 KVRIVAS 162  
Db 12 KVRIVAS 18

RESULT 5  
Q9HVX4 PRELIMINARY; PRT; 148 AA.  
ID Q9HVX4  
AC Q9HVX4:  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA4441.  
GN PA4441.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody R.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL: AE004858; AAG07829.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 148 AA; 16418 MW; 4DD9E4B217EDD2B8 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 ARLPNA 116  
Db 26 ARLPNA 32

RESULT 6  
Q9JYB9 PRELIMINARY; PRT; 197 AA.  
ID Q9JYB9  
AC Q9JYB9:  
DT 01-OCT-2000 (TReMBLrel. 15, Created)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:34:30 : Search time 79.75 Seconds  
(without alignments)  
160.397 Million cell updates/sec

Title: US-09-668-021-12  
Perfect score: 211  
Sequence: 1 MOPSLAPICLICLVHAAFCV.....KPRPGARCAKANAQALENAV 211

Scoring table:  
Gapop 60.0, Gapext 60.0

Searched: 412676 seqs, 60623988 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database:

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	100.0	211	21	Murine TGF-beta bi
2	108	51.2	213	21	Rat TGF-beta bindi
3	32	15.2	176	21	Bovine TGF-beta bi
4	32	15.2	213	21	Human DAN/Cerberus
5	32	15.2	213	21	Human TGF-beta bin
6	32	15.2	213	21	Human TGF-beta bin
7	32	15.2	213	21	Human TGF-beta bin
8	32	15.2	213	21	Human TGF-beta bin
9	32	15.2	213	21	Human secreted pro
10	32	15.2	367	21	Human DAN/Cerberus
11	10	4.7	50	20	Human 5' EST secre

12	8	3.8	428	20	AAV52704	Human ELK-1 protei
13	8	3.8	428	20	AAV32824	Human ELK-1. Homo
14	8	3.8	481	21	AA634091	Zea mays protein f
15	8	3.8	490	21	AA634090	Zea mays protein f
16	8	3.8	512	21	AA634089	Zea mays protein f
17	7	3.3	23	15	AA657116	Immunostimulatory
18	7	3.3	86	15	AA67202	Immunostimulatory
19	7	3.3	99	21	AA68229	Guinea pig beta-2-
20	7	3.3	99	21	AA68283	Guinea pig beta-2-
21	7	3.3	99	22	AA68644	Guinea pig beta 2m
22	7	3.3	201	12	AA610425	E. coli lpp gene pr
23	7	3.3	219	21	AA642285	Arabidopsis thalia
24	7	3.3	265	21	AA642284	Arabidopsis thalia
25	7	3.3	282	21	AA642283	Arabidopsis thalia
26	7	3.3	431	21	AA647031	Arabidopsis thalia
27	7	3.3	500	21	AA647030	Arabidopsis thalia
28	7	3.3	511	21	AA647029	Arabidopsis thalia
29	6	2.8	10	12	AA610268	Human ventricular
30	6	2.8	10	18	AA640709	Peptide which bind
31	6	2.8	10	18	AA640710	Peptide which bind
32	6	2.8	10	18	AA640711	Peptide which bind
33	6	2.8	10	18	AA640712	Peptide which bind
34	6	2.8	12	6	AA650474	Determinant site o
35	6	2.8	12	10	AA692004	Human prolactin C-
36	6	2.8	13	7	AA660742	Sequence from synt
37	6	2.8	14	17	AA695155	bcl-x(L)/bcl-2 ass
38	6	2.8	15	19	AA645818	Peptide recognised
39	6	2.8	15	19	AA645613	Peptide recognised
40	6	2.8	20	13	AA627052	N-terminal prolact
41	6	2.8	22	22	AA687425	Human gene 4 encod
42	6	2.8	41	20	AA613116	Human secreted pro
43	6	2.8	43	17	AA693501	E99, monoclonal an
44	6	2.8	44	19	AA669242	Clone BV66_1 prote
45	6	2.8	44	20	AA674042	Human D2H binding
46	6	2.8	45	21	AA638165	Human secreted pro
47	6	2.8	46	20	AA612088	Human 5' EST secre
48	6	2.8	46	21	AA654384	Human pancreatic c
49	6	2.8	47	17	AA693140	Actinomadura sp. F
50	6	2.8	48	21	AA608651	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	
AAV96432	standard; Protein: 211 AA.
ID	AAV96432
XX	
AC	AAV96432;
XX	
DT	12-SEP-2000 (first entry)
XX	
DE	Murine TGF-beta binding protein (BEER).
XX	
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX	
OS	Mus musculus.
XX	
PN	W0200032773-A1.
PD	08-JUN-2000.
XX	
PF	24-NOV-1999; 99WC-US27990.
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
XX	
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
XX	Van Ness J, Winkler DG;
DR	WPI: 2000-412321/35.

DR N-PSDB; AAA29058.  
 XX Nucleic acids (I) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 XX  
 PS Claim 5; Page 124; 162pp; English.  
 CC This shows a murine transforming growth factor-beta (TGF-beta)  
 CC binding protein designated mBEER. The cdna and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BEER expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cdna or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BEER by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BEER polypeptides. The nucleic acids may be used for recombinant  
 CC production of BEER, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BEER may be used to raise  
 CC antibodies and for identification of BEER modulators. BEER antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 XX  
 SQ Sequence 211 AA;

Query Match 100.0%; Score 211; DB 21; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-210; Mismatches 0; Gaps 0;

Matches 211; Conservative 0; Indels 0; Gaps 0;  
 QY 1 MOPSTAPCLICLVAHAFCAVESGOGMOAFRNDATFVPGTGEPEPPENNOTMNAENG 60  
 DB 1 mgpslapclcllhaatcaveggqgafndatevlpjlgyppeppennqtmnaeng 60  
 QY 61 GRPHNHPYDAKDVSEYSCRELHTRRLTLDGPCRSAPKVTLELVSGGCGPARLLPNAIGRV 120  
 DB 61 grphnhydakdvseysecretlhytrrltldgpcrsakvtelevsgcgparllpnaigrv 120  
 QY 121 KWMRNGPDPFCIPDRRAOVOLLCGGAAPSRKVRILVASCCKRLTRFHNQSELKDF 180  
 DB 121 kwmrnpdpfcipdryraqvqlldpggaaprsrkvrilvasckckrltrfhnqselkdf 180  
 QY 181 GETARPGKRRPRPGARGAKANOAELENAY 211  
 DB 181 gpetarpgkrrkprpgargakanaeleenay 211

## RESULT 2

AAV96433  
 ID AAY96433 standard; Protein: 213 AA.

AC AAY96433;  
 XX  
 DT 12-SEP-2000 (first entry)  
 XX  
 DE Rat TGF-beta binding protein (BEER).  
 XX  
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
 OS Rattus norvegicus.  
 XX  
 PN WO200032773-A1.  
 XX  
 PD 08-JUN-2000.  
 XX  
 PF 24-NOV-1999; 99WO-US27990.  
 XX  
 PR 27-NOV-1998; 98US-0110283.  
 XX  
 PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
 PI Van Ness J, Winkler DG;  
 XX  
 DR WPI: 2000-412321/35.  
 DR N-PSDB; AAA29059.  
 XX  
 PT Nucleic acids (I) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 XX  
 PS Claim 6; Page 125-126; 162pp; English.

CC This shows a rat transforming growth factor-beta (TGF-beta) binding  
 CC protein designated rBEER. The cdna and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BEER expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cdna or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BEER by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BEER polypeptides. The nucleic acids may be used for recombinant  
 CC production of BEER, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BEER may be used to raise  
 CC antibodies and for identification of BEER modulators. BEER antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 XX  
 SQ Sequence 213 AA;

Query Match 51.2%; Score 108; DB 21; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-104; Mismatches 0; Gaps 0;

Matches 108; Conservative 0; Indels 0; Gaps 0;  
 QY 88 TDGPCRSAPKVTLELVSGGCGPARLLPNAIGRVKWMRNGPDPFCIPDRRAOVOLLC 147  
 DB 90 tdgpcrsakvtelevsgcgparllpnaigrvkwmrnpdpfcipdryraqvqlldp 149  
 QY 148 GAAPSRKVRILVASCCKRLTRFHNQSELKDFGETRPOKGRKPR 195  
 DB 150 ggaaprsrkvrilvasckckrltrfhnqselkdfgetarpgkrrkpr 197

## RESULT 3

AAV96434  
 ID AAY96434 standard; Protein: 176 AA.

AC AAY96434;  
 XX  
 DT 12-SEP-2000 (first entry)  
 XX  
 DE Bovine TGF-beta binding protein (BEER).  
 XX  
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
 OS Bos taurus.  
 XX  
 PN WO200032773-A1.  
 XX  
 PD 08-JUN-2000.  
 XX  
 PF 24-NOV-1999; 99WO-US27990.  
 XX  
 PR 27-NOV-1998; 98US-0110283.  
 XX  
 PA (DARW-) DARWIN DISCOVERY LTD.  
 XX  
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
 PI Van Ness J, Winkler DG;  
 DR WPI: 2000-412321/35.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:35:24 : Search time 47.66 Seconds  
(without alignments)  
99,626 Million cell updates/sec

Title: US-09-668-021-12

Perfect score: 211  
Sequence: 1 MOPSLAPFLICLVHAARCA.....KPRPGARCAKANOAELENAV 211

Scoring table:

Gapop 60.0, Gapext 60.0

arched: 212252 seqs, 22503292 residues

Word size: 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database:

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.3	99	2	US-08-484-905-61
2	7	3.3	99	3	US-08-481-985B-61
3	7	3.3	99	4	US-08-370-476-61
4	6	2.8	10	1	US-08-250-789A-94
5	6	2.8	10	1	US-08-250-789A-95
6	6	2.8	10	1	US-08-250-789A-96
7	6	2.8	10	1	US-08-250-789A-97
8	6	2.8	14	1	US-08-333-565-18
9	6	2.8	14	2	US-08-661-479-18
10	6	2.8	24	1	US-08-220-272A-6
11	6	2.8	41	4	US-08-905-223-400
12	6	2.8	43	2	US-08-488-161-68
13	6	2.8	43	3	US-09-273-685-68
14	6	2.8	43	5	PCT-US95-11934-68
15	6	2.8	47	2	US-08-282-197C-44
16	6	2.8	68	2	US-08-836-791-5
17	6	2.8	97	2	US-08-245-511-30
18	6	2.8	97	2	US-08-600-993A-30
19	6	2.8	125	3	US-08-985-526-25
20	6	2.8	128	1	US-08-666-798-2
21	6	2.8	128	1	US-08-892-692-2
22	6	2.8	128	2	US-09-096-071-2
23	6	2.8	167	2	US-08-282-197C-2
24	6	2.8	166	4	US-09-475-316A-13
25	6	2.8	199	3	US-08-737-248-7
26	6	2.8	204	1	US-08-333-565-2
27	6	2.8	204	2	US-08-661-479-2

28	6	2.8	204	2	US-08-733-505A-1	Sequence 1, Appli
29	6	2.8	204	2	US-08-733-505A-12	Sequence 12, Appl
30	6	2.8	204	2	US-08-733-505A-13	Sequence 13, Appl
31	6	2.8	204	2	US-08-733-505A-14	Sequence 14, Appl
32	6	2.8	204	2	US-08-717-123-3	Sequence 3, Appli
33	6	2.8	207	4	US-08-652-877-13	Sequence 13, Appl
34	6	2.8	207	4	US-08-476-515A-13	Sequence 13, Appl
35	6	2.8	216	2	US-08-821-637-3	Sequence 3, Appl
36	6	2.8	224	3	US-08-857-534-10	Sequence 10, Appl
37	6	2.8	224	5	PCT-US95-04971-10	Sequence 10, Appl
38	6	2.8	253	3	US-08-985-526-27	Sequence 27, Appl
39	6	2.8	256	2	US-08-727-688-33	Sequence 33, Appl
40	6	2.8	300	4	US-09-433-898-2	Sequence 2, Appli
41	6	2.8	330	4	US-09-188-930-144	Sequence 144, App
42	6	2.8	330	4	US-09-188-930-278	Sequence 278, App
43	6	2.8	343	2	US-08-933-750C-13	Sequence 13, Appl
44	6	2.8	343	4	US-09-234-613-13	Sequence 13, Appl
45	6	2.8	351	1	US-08-196-350-1	Sequence 1, Appli
46	6	2.8	388	1	US-08-087-772A-2	Sequence 2, Appli
47	6	2.8	395	3	US-08-981-825-6	Sequence 6, Appli
48	6	2.8	395	4	US-09-480-784-6	Sequence 6, Appli
49	6	2.8	400	1	US-07-916-901-6	Sequence 6, Appli
50	6	2.8	400	1	US-07-783-602C-1	Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-08-484-905-61  
Sequence 61, Application US/08484905  
Patent No. 5976551  
GENERAL INFORMATION:  
APPLICANT: Motiez, Estelle  
APPLICANT: Apostado, Jean-Pierre  
TITLE OF INVENTION: An Altered Major Histocompatibility  
TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the  
NUMBER OF SEQUENCES: 127  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS-/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,905  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/801,818  
FILING DATE: 05-DEC-1991  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/792,473  
FILING DATE: 15-NOV-1991  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 03495.0106-03000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:

LENGTH: 99 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-484-905-61

Query Match 3.3%; Score 7; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLVHAAF 18  
|||||  
DB 64 LLVHAAF 70

## RESULT 2

US-08-481-985B-61  
Sequence 61, Application US/08481985B  
Patent No. 6011146  
GENERAL INFORMATION:  
APPLICANT: Motiez, Estelle  
APPLICANT: Abastado, Jean-Pierre  
APPLICANT: Kourilsky, Philippe  
TITLE OF INVENTION: Altered Major Histocompatibility Complex  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunnet  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,985B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/801,818  
FILING DATE: 05-DEC-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/792,473  
FILING DATE: 15-NOV-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03495, 0106-04000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-481-985B-61

Query Match 3.3%; Score 7; DB 3; Length 99;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLVHAAF 18  
|||||

DB 64 LLVHAAF 70

## RESULT 3

US-08-370-476-61  
Sequence 61, Application US/08370476  
Patent No. 6153408  
GENERAL INFORMATION:  
APPLICANT: Motiez, Estelle  
APPLICANT: Abastado, Jean-Pierre  
APPLICANT: Kourilsky, Philippe  
APPLICANT: Lone, Yu-Chun  
APPLICANT: Ojcius, David  
APPLICANT: Carrouge, Amanda  
TITLE OF INVENTION: Altered Major Histocompatibility Complex  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 127  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunnet  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/370,476  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/117,575  
FILING DATE: 07-SEP-1993  
APPLICATION NUMBER: US 08/072,787  
FILING DATE: 06-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/801,818  
FILING DATE: 05-DEC-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/792,473  
FILING DATE: 15-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 05243, 0001-01000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-370-476-61

Query Match 3.3%; Score 7; DB 4; Length 99;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLVHAAF 18  
|||||  
DB 64 LLVHAAF 70

RESULT 4  
US-08-250-789A-94  
Sequence 94, Application US/08250789A  
Patent No. 5635397



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:36:29 ; Search time 54.37 Seconds  
(without alignments)  
298.422 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 213

Sequence: 1 MOLSLAPCLACLIVHAAFAV.....KPRPRARAKAQAQALENAY 213

Scoring table:

Gapop 60.0 , Gapext 60.0

Archived: 219241 seqs, 76174552 residues

rd size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database: PIR.68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	428	1	transforming prote
2	8	3.8	429	2	elk1 protein - mou
3	8	3.8	453	1	ubiquinol-cytochr
4	8	3.8	850	2	EDB kinase activa
5	7	3.3	99	1	beta-2-microglobul
6	7	3.3	130	2	hypothetical prote
7	7	3.3	148	2	hypothetical prote
8	7	3.3	197	2	hypothetical prote
9	7	3.3	197	2	hypothetical prote
10	7	3.3	268	2	hypothetical prote
11	7	3.3	283	2	hypothetical prote
12	7	3.3	321	2	hypothetical prote
13	7	3.3	327	2	conserved hypotet
14	7	3.3	327	2	hypothetical prote
15	7	3.3	328	2	probable LACI-type
16	7	3.3	354	2	hypothetical prote
17	7	3.3	398	1	poly(epsilon-doxin
18	7	3.3	459	2	poly(epsilon-doxin
19	7	3.3	461	2	ID-myo-inositol-tr
20	7	3.3	481	2	lincomycin resista
21	7	3.3	531	2	tyrosine decarboxy
22	7	3.3	542	2	hypothetical prote
23	7	3.3	547	2	collagen alpha cha
24	7	3.3	562	2	hypothetical prote
25	7	3.3	628	2	glucan 1,3-beta-gl
26	7	3.3	632	2	surface glycoprote
27	7	3.3	654	2	beta-glucosidase (
28	7	3.3	681	2	outer membrane hem
29	7	3.3	747	2	hypothetical prote

30	7	3.3	821	2	probable Na/H anti
31	7	3.3	829	2	translation initia
32	7	3.3	900	2	probable InbB - My
33	7	3.3	962	2	translation initia
34	7	3.3	962	2	translation initia
35	7	3.3	1168	1	myosin heavy chain
36	7	3.3	1308	2	hypothetical prote
37	7	3.3	1940	2	myosin heavy chain
38	6	2.8	22	2	ribosomal protein
39	6	2.8	49	2	neurotoxin Tx2 - s
40	6	2.8	63	2	hypothetical prote
41	6	2.8	65	2	malx protein - Kle
42	6	2.8	89	2	hypothetical prote
43	6	2.8	100	2	hypothetical prote
44	6	2.8	102	2	unknown protein en
45	6	2.8	103	2	hypothetical prote
46	6	2.8	108	1	thioredoxin C-2 -
47	6	2.8	109	2	ribosomal protein
48	6	2.8	110	2	50S ribosomal subu
49	6	2.8	110	2	ribosomal protein
50	6	2.8	111	2	ribosomal protein

#### ALIGNMENTS

RESULT 1  
TVHUEK  
transforming protein elk-1 - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Mar-1993 #sequence-revision 31-Mar-1993 #text-change 18-Jun-1999  
C/Accession: A41354; S54721  
R: Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.  
Science 244, 66-70, 1989  
A>Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near transloc  
A/Reference number: A41354; MUID:89203250  
A/Accession: A41354  
A/Molecule type: mRNA  
A/Residues: 1-428 <RAO>  
A/Cross-references: GB:M25269; NID:9538208; PIDN:AAA52384.1; PID:9538209  
R:Gille, H.; Kortentjan, M.; Thoma, O.; Moosaw, C.; Slaughter, C.; Cobb, M.H.; Shaw, EMO J. 14, 951-962, 1995  
A>Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and  
A/Reference number: S54721; MUID:95196758  
A/Accession: S54721  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 318-328, 'XX', '331, '336-364; '380-388, 'X', '390-392, 'X', '394-400, 'XX', '403-405, 'X'  
C/Genes: GDB:ELK1  
A/Cross-references: GDB:119867; OMIM:311040  
A/Map position: Xp11.2-Xp11.2  
C/Superfamily: elk-1-transforming protein; ets DNA-binding domain homology  
C/Keywords: DNA binding; oncogene; transforming protein  
F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 3.8%; Score 8; DB 1; Length 428;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKRRPR 196  
Db 310 POKRRPR 317

#### RESULT 2

JC4965  
elk1 protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Dec-1996 #sequence-revision 31-Dec-1996 #text-change 16-Jul-1999  
C/Accession: JC4965; I48339; I48340; S54908  
R:Greyn, D.; Ung, S.; Denhez, F.; Delhem, M.; Quatannens, B.; Begue, A.; Stehelin, D.

Gene 174, 185-188, 1996  
A:Title: Structure and organization of the mouse elk1 gene.

A:Reference number: J04965; MUID:97017146

A:Accession: J04965

A:Molecule type: mRNA

A:Residues: 1-429 <RES>

A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635

A:Experimental source: embryo

R:Giovane, A.; Pintzas, A.; Maira, S.M.; Sobieszczuk, P.; Wasyluk, B.

Genes Dev. 8, 1502-1513, 1994

A:Title: Net, a new ets transcription factor that is activated by Ras.

A:Reference number: A53837; MUID:95047310

A:Accession: I48339

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-429 <RES>

A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635

A:Accession: I48340

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 5-132, 134-224 <RES>

A:Cross-references: EMBL:Z6939; NID:g535922; PIDN:CAA85391.1; PID:g535923

A:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which h

C:Genetics:

A:Gene: elk1

A:Introns: 70/3; 219/3; 363/3; 397/3

C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology

F:7-86/Domain: ets DNA-binding domain homology <ETS>

#### Query Match

Best Local Similarity 3.8%; Score 8; DB 2; Length 429;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 POKGRKPR 196

|||||||

Db 311 POKGRKPR 318

#### RESULT 3

A32629

ubiquinol--cytochrome-c reductase (EC 1.10.2.2) core protein II precursor - human

C:Species: Homo sapiens (man)

C>Date: 21-May-1990 #sequence\_revision 14-Jul-1994 #text\_change 16-Jul-1999

C:Accession: A32629

R:Hosokawa, Y.; Suzuki, H.; Toda, H.; Nishikimi, M.; Ozawa, T.

J. Biol. Chem. 264, 13483-13488, 1989

A:Title: Complementary DNA encoding core protein II of human mitochondrial cytochrome bc

A:Reference number: A32629; MUID:89340421

A:Accession: A32629

A:Molecule type: mRNA

A:Residues: 1-453 <HOS>

A:Cross-references: GB:J04973; NID:g180927; PIDN:AAA35710.1; PID:g180928

A>Note: the authors translated the codon AGA for residue 360 as Thr

C:Genetics:

A:Gene: GDB:UDCR2

A:Cross-references: GDB:141851; OMIM:191329

A:Map position: 16p12.3-16p12.3

C:Superfamily: mitochondrial processing peptidase alpha chain

C:Keywords: hydrolase; mitochondrial matrix; mitochondrion; oxidative phosphorylation; c

F:1-14/Domain: transit peptide (mitochondrion) #status predicted <TMP>

F:15-453/Product: ubiquinol--cytochrome-c reductase core protein II #status predicted <N

#### Query Match

Best Local Similarity 3.8%; Score 8; DB 1; Length 453;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LVHAAFYA 20

|||||||

Db 266 LVHAAFYA 273

#### RESULT 4

J05700

ERB kinase activator alpha, brain and thymus - human

C:Species: Homo sapiens (man)

C>Date: 25-Nov-1997 #sequence\_revision 25-Nov-1997 #text\_change 21-Jul-2000

C:Accession: J05700

R:Higashiyama, S.; Horikawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.; Mi

J. Biochem. 122, 675-680, 1997

A:Title: A novel brain-derived member of the epidermal growth factor family that inte

A:Reference number: J05700; MUID:98006324

A:Accession: J05700

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-850 <HIG>

A:Cross-references: DBJ:AB005060; NID:g2626738; PIDN:BAA23417.1; PID:g2626739

A:Experimental source: SK-NSH cell

A:Comment: This protein is a member of the epidermal growth factor family. It is func

ating the differentiation of MDA-MB-453 cells.

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: glycoprotein

F:258-311/Domain: Ig-like #status predicted <IGL>

F:345-381/Domain: EGF homology <EGF>

F:346-381/Domain: EGF-like #status predicted <EGF2>

F:147,278,451/Binding site: carbohydrate (Asn) (covalent) #status predicted

#### Query Match

Best Local Similarity 3.8%; Score 8; DB 2; Length 850;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 PRPRARGA 202

|||||||

Db 725 PRPRARGA 732

#### RESULT 5

M0GPR2

beta-2-microglobulin - guinea pig

N:Alternate names: class I histocompatibility antigen beta chain

C:Species: Cavia porcellus (guinea pig)

C>Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 31-Jan-1997

C:Accession: A02181

R:Wolfe, P.B.; Gebra, J.J.

Mol. Immunol. 17, 1493-1505, 1980

A:Title: The primary structure of guinea pig beta2-microglobulin.

A:Reference number: A02181; MUID:82057805

A:Accession: A02181

A:Molecule type: protein

A:Residues: 1-99 <WOL>

C:Complex: heterodimer with MHC class I histocompatibility antigen alpha chain

C:Function:

A:Description: necessary for the expression of MHC class I histocompatibility antigen

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: heterodimer

F:18-82/Domain: immunoglobulin homology <IMN>

F:25-80/Disulfide bonds: #status predicted

#### Query Match

Best Local Similarity 3.3%; Score 7; DB 1; Length 99;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LTVHAF 18

|||||||

Db 64 LTVHAF 70

#### RESULT 6

D83305

hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: D83305

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:57 ; Search time 31.03 Seconds

(without alignments)  
235,141 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 213

Sequence: 1 MQLSLAPCLACLHAAFAVA.....KPRPRAGAKAQAELNNAY 213

Scoring table:

Gapop 60.0 , Gapext 60.0

riched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	428	1	ELK1_HUMAN
2	8	3.8	429	1	ELK1_MOUSE
3	8	3.8	453	1	UCR2_HUMAN
4	8	3.8	850	1	NRG2_HUMAN
5	7	3.3	97	1	VE7_HPV4
6	7	3.3	99	1	B2M6_CAVPO
7	7	3.3	110	1	RL22_LEPIN
8	7	3.3	327	1	YPHF_ECOLI
9	7	3.3	459	1	IP3K_RAT
10	7	3.3	461	1	IP3K_HUMAN
11	7	3.3	481	1	LMRA_STRLN
12	7	3.3	531	1	TYD2_PAPSO
13	7	3.3	547	1	CAP1_EPRHMT
14	7	3.3	632	1	VE16_STYV
15	7	3.3	829	1	IF2_HABIN
16	7	3.3	900	1	IF2_MYCTU
17	7	3.3	924	1	IF2_MYCLE
18	7	3.3	1168	1	MYSC_ACACA
19	6	2.8	49	1	TX25_PHONI
20	6	2.8	53	1	LHAI_RHOAC
21	6	2.8	65	1	MAIX_KLEPN
22	6	2.8	100	1	VG10_HSVB
23	6	2.8	107	1	TH12_HSVB
24	6	2.8	110	1	RL22_ECOLI
25	6	2.8	111	1	RL22_ACHLA
26	6	2.8	112	1	RL22_SPICI
27	6	2.8	113	1	RL22_THERT
28	6	2.8	114	1	NL21_TOBAC
29	6	2.8	115	1	TIAF_HUMAN
30	6	2.8	115	1	TIAF_MOUSE
31	6	2.8	118	1	REV_HVILW
32	6	2.8	131	1	RL22_PHYSI
33	6	2.8	146	1	HBBI_IGUIG

34	6	2.8	148	1	RK22_MAIZE	P06589 zea mays (m
35	6	2.8	149	1	RK22_ORYSA	P12140 oryza sativ
36	6	2.8	167	1	B3AR_MERUN	O70433 meriones un
37	6	2.8	168	1	YWY1_CAEEL	Q11088 caenorhabdi
38	6	2.8	182	1	YCVO_YEAST	P25654 saccharomyc
39	6	2.8	193	1	HS72_CANAL	P46587 candida alb
40	6	2.8	196	1	REPB_CHICK	P41263 gallus gall
41	6	2.8	210	1	TRPF_KUDILA	P13997 kiuyvetromyc
42	6	2.8	213	1	PYRE_HAEIN	P43855 haemophilus
43	6	2.8	215	1	DER1_DERPT	P49273 dermatophag
44	6	2.8	216	1	FGFR_HUMAN	O60258 homo sapien
45	6	2.8	216	1	FGFR_MOUSE	O70627 mus musculu
46	6	2.8	223	1	GRPM_MOUSE	O88396 mus musculu
47	6	2.8	227	1	PRL_HUMAN	P01236 homo sapien
48	6	2.8	227	1	PRL_MACMU	P55151 macaca mula
49	6	2.8	237	1	YJGT_ECOLI	P39333 escherichia
50	6	2.8	261	1	DHBB_RICPR	O92eal ticketstia

#### ALIGNMENTS

```

RESULT 1
ID      ELK1_HUMAN      STANDARD:      PRT:      428 AA.
AC      P19419; O75606; Q9UDM4; O95058;
DT      01-NOV-1990 (Rel. 16, Created)
DT      01-OCT-2000 (Rel. 40, Last sequence update)
DT      01-OCT-2000 (Rel. 40, Last annotation update)
DE      ETS-DOMAIN PROTEIN ELK-1.
GN      ELK1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OA      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89203250; PubMed=2539641;
RA      Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,
RT      Reddy E.S.P.;
RL      "elk, tissue-specific ets-related genes on chromosomes X and 14 near
       translocation breakpoints.";
       Science 244:66-70(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99013876; PubMed=9795224;
RA      Harindranath N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;
RT      "The human elk-1 gene family: the functional gene and two processed
       pseudogenes embedded in the Igh locus.";
       Gene 221:215-224(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Grahnam D.;
RL      Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A. (ISOFORM 2).
RA      Aryee D.N.T., Kovar H.;
RT      "Novel family members Huer71, ELFR, and ELKV among ETS-related genes
       coexpressed with EMS-FLI1 in Ewing tumor cell lines.";
       Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      DOMAINS.
RX      MEDLINE=92334979; PubMed=1630903;
RA      Jankecht R., Nordheim A.;
RT      "Elk-1 protein domains required for direct and SRF-assisted
       DNA-binding.";
       Nucleic Acids Res. 20:3317-3324(1992).
CC      -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
       SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
       ELEMENT AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
       ELEMENT.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR.
CC      -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE), AND 2/ELKV; ARE

```

```

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: LUNG AND TESTIS.
CC -1- PFM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,
CC AND ERK).
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25269; AA52384.1; -
DR EMBL: AF080616; AAC82466.1; -
DR EMBL: AL009172; CAI15659.1; -
DR EMBL: AF000672; AAD00862.1; -
DR PIR: A41354; TVHUEK.
DR HSSP: P14921; 2STW.
DR TRANSFAC: T00250; -
DR MIM: 311040; -
DR Interpro: IPR000418; -
DR Pfam: PF00178; Ets.1.
DR PRINTS: PR00454; ETSDOMAIN.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
DR Phosphorylation; Alternative splicing.
FT DNA_BIND 5 86
FT VARSPIC 91 95 VAGGS -> SHCAP (IN ISOFORM 2).
FT VARSPLIC 96 428 MISSING (IN ISOFORM 2).
FT CONFLICT 183 S -> N (IN REF. 1)
FT SEQUENCE 428 AA; 4488 MW; 68F71F8DB9D38CA CRC64;

Query Match 3.8%; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKGRKPR 196
DB 310 POKGRKPR 317

RESULT 2
ELK1_MOUSE
ID ELK1_MOUSE STANDARD; PRT; 429 AA.
AC P41669;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ETS-DOMAIN PROTEIN ELK-1.
GN ELK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=97017146; PubMed=8663747;
RA Grevin D., Ung S., Denhez F., Dehem M., Quatnenns B., Begue A.,
RA Stelien D., Martin P.;
RT "Structure and organization of the mouse elk1 gene.";
RL Gene 174:185-188(1996).
RN [2]
RP SEQUENCE OF 5-224 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95047310; PubMed=7958835;
RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Wasylyk B.,
RA "Net, a new ets transcription factor that is activated by Ras.";
RT

```

```

RL Genes Dev. 8:1502-1513(1994).
CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
CC ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87257; CAA60715.1; -
DR EMBL: Z36939; CAA85391.1; -
DR HSSP: P14921; 2STW.
DR MGD: MGI:101833; ELK1.
DR Interpro: IPR000418; -
DR Pfam: PF00178; Ets.1.
DR PRINTS: PR00454; ETSDOMAIN.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
DR Phosphorylation.
FT DNA_BIND 5 86
FT CONFLICT 133 P -> T (IN REF. 2).
FT SEQUENCE 429 AA; 45243 MW; B615B977731D54F CRC64;

Query Match 3.8%; Score 8; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 POKGRKPR 196
DB 311 POKGRKPR 318

RESULT 3
UCR2_HUMAN
ID UCR2_HUMAN STANDARD; PRT; 453 AA.
AC P22695;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UBIQUITIN-L-CYTOCHROME C REDUCTASE COMPLEX CORE PROTEIN 2 PRECURSOR
DE (EC 1.10.2.2) (COMPLEX III SUBUNIT II).
GN UQCRC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89340421; PubMed=2547763;
RA Hosokawa Y., Suzuki H., Toda H., Nishikimi M., Ozawa T.;
RT "Complementary DNA encoding core protein II of human mitochondrial
RT cytochrome bc1 complex. Substantial diversity in deduced primary
RT structure from its yeast counterpart.";
RL J. Biol. Chem. 264:13483-13488(1989).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C
CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH
CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THE CORE PROTEIN
CC 2 IS REQUIRED FOR THE ASSEMBLY OF THE COMPLEX.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q +
CC 2 FERROCYTOCHROME C.
CC -1- SUBUNIT: BCI COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS
CC

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:20 ; Search time 89.8 Seconds  
(without alignments)  
313.819 Million cell updates/sec

Title: US-09-668-021-14  
Perfect score: 213  
Sequence: 1 MQLSLAPCLACLVHAAFA.....KPPRRARAKANQAELENAY 213

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

rchd: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

SPTREMBL\_16:\*  
1: sp.\_archaea:\*  
2: sp.\_bacteria:\*  
3: sp.\_fungi:\*  
4: sp.\_human:\*  
5: sp.\_invertebrate:\*  
6: sp.\_mammal:\*  
7: sp.\_mhc:\*  
8: sp.\_organelle:\*  
9: sp.\_phage:\*  
10: sp.\_plant:\*  
11: sp.\_rodent:\*  
12: sp.\_unclassified:\*  
13: sp.\_vertebrate:\*  
14: sp.\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	276	13	Q9YGD6
2	8	3.8	276	13	Q9PT38
3	8	3.8	276	13	Q9PT37
4	8	3.8	276	13	Q9PT36
5	8	3.8	499	2	Q9RDF5
6	7	3.3	89	10	Q9FH55
7	7	3.3	109	5	Q9GTJ4
8	7	3.3	125	11	Q9E001
9	7	3.3	130	2	Q9I0C0
10	7	3.3	148	2	Q9HYX4
11	7	3.3	150	2	Q9E0U6
12	7	3.3	151	2	Q9EV14
13	7	3.3	151	2	Q9E0Z2
14	7	3.3	151	2	Q9E0Z1
15	7	3.3	151	2	Q9E0Z0
16	7	3.3	151	2	Q9E0Y9
17	7	3.3	151	2	Q9E0Y8
18	7	3.3	151	2	Q9E0Y6
19	7	3.3	151	2	Q9E0Y5

20	7	3.3	151	2	Q9E0Y4	Q9E0Y4 haemophilus
21	7	3.3	151	2	Q9E0Y3	Q9E0Y3 haemophilus
22	7	3.3	151	2	Q9E0Y2	Q9E0Y2 haemophilus
23	7	3.3	151	2	Q9E0Y1	Q9E0Y1 haemophilus
24	7	3.3	151	2	Q9E0Y0	Q9E0Y0 haemophilus
25	7	3.3	151	2	Q9E0X9	Q9E0X9 haemophilus
26	7	3.3	151	2	Q9E0X8	Q9E0X8 haemophilus
27	7	3.3	151	2	Q9E0X7	Q9E0X7 haemophilus
28	7	3.3	151	2	Q9E0X6	Q9E0X6 haemophilus
29	7	3.3	151	2	Q9E0X5	Q9E0X5 haemophilus
30	7	3.3	151	2	Q9E0X4	Q9E0X4 haemophilus
31	7	3.3	151	2	Q9E0X3	Q9E0X3 haemophilus
32	7	3.3	151	2	Q9E0X2	Q9E0X2 haemophilus
33	7	3.3	151	2	Q9E0X1	Q9E0X1 haemophilus
34	7	3.3	151	2	Q9E0U0	Q9E0U0 haemophilus
35	7	3.3	151	2	Q9E0U7	Q9E0U7 pasteurilla
36	7	3.3	151	2	Q9E0D1	Q9E0D1 haemophilus
37	7	3.3	151	2	Q9E0D0	Q9E0D0 haemophilus
38	7	3.3	151	2	Q9E0C9	Q9E0C9 haemophilus
39	7	3.3	151	2	Q9E0C8	Q9E0C8 haemophilus
40	7	3.3	151	2	Q9E0C2	Q9E0C2 actinobacil
41	7	3.3	151	2	Q9E0Z0	Q9E0Z0 haemophilus
42	7	3.3	151	2	Q9E0T0	Q9E0T0 haemophilus
43	7	3.3	151	2	Q9E0T0	Q9E0T0 haemophilus
44	7	3.3	151	2	Q9E0T9	Q9E0T9 haemophilus
45	7	3.3	186	10	Q9SDR7	Q9SDR7 forsythia x
46	7	3.3	197	2	Q9JYB9	Q9JYB9 neisseria m
47	7	3.3	197	2	Q9JYB9	Q9JYB9 neisseria m
48	7	3.3	239	14	Q9IBJ1	Q9IBJ1 turkey herp
49	7	3.3	251	2	Q9I1N0	Q9I1N0 streptomyces
50	7	3.3	268	2	P94983	P94983 mycobacteri

## ALIGNMENTS

RESULT 1  
Q9YGD6 PRELIMINARY; PRT; 276 AA.  
AC Q9YGD6; 01-MAY-1999 (TREMURel. 10, Created)  
DT 01-MAY-1999 (TREMURel. 10, Last sequence update)  
DE 01-MAR-2001 (TREMURel. 16, Last annotation update)  
DE CARBOXYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Euteleostei; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Proctacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99185307; PubMed=10082666;  
RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
RT "Cloning and expression of two carbonyl reductase-like 20beta-hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow trout (Oncorhynchus mykiss).";  
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).  
CC -I- FAMILY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
CC EMBL: AF100930; AAD20217.1; -.  
DR HSSP; P50163; ZAE1.  
DR InterPro: IPR002347; -.  
DR InterPro: IPR002347; -.  
DR Pfam; PF0106; adh\_short; 1.  
DR PRINTS; PR00081; GDRDH.  
DR PRINTS; PR00080; SDRFAMILY.  
KW Oxidoreductase.  
SQ SEQUENCE 276 AA; 30111 MW; 135604A38D2940CD CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 28 AFKNDATE 35  
 |||||  
 Db 93 AFKNDATE 100

## RESULT 2

O9PT38 PRELIMINARY; PRT; 276 AA.

ID O9PT38; PRELIMINARY; PRT; 276 AA.  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_Taxid=8022;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Blochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 FAMILY.  
 DR EMBL: AF100933; AAD2092.1; -.  
 DR HSSP: P50163; 2AEL.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDHRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SQ SEQUENCE 276 AA; 30239 MW; 134B52054C751C28 CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATE 35  
 |||||  
 Db 93 AFKNDATE 100

## RESULT 3

O9PT37 PRELIMINARY; PRT; 276 AA.

ID O9PT37; PRELIMINARY; PRT; 276 AA.  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE B.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_Taxid=8022;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Blochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 FAMILY.  
 DR EMBL: AF100932; AAD2091.1; -.

DR HSSP: P50163; 2AEL.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDHRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SQ SEQUENCE 276 AA; 30127 MW; 1350E4211D34225D CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATE 35  
 |||||  
 Db 93 AFKNDATE 100

## RESULT 4

O9PT36 PRELIMINARY; PRT; 276 AA.

ID O9PT36; PRELIMINARY; PRT; 276 AA.  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 16, Last annotation update)  
 DE CARBONYL REDUCTASE/20BETA-HYDROXYSTEROID DEHYDROGENASE A.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_Taxid=8022;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9185307; PubMed=10082666;  
 RA Guan G.J., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;  
 RT "Cloning and expression of two carbonyl reductase-like 20beta-  
 hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow  
 trout (Oncorhynchus mykiss).";  
 RL Blochem. Biophys. Res. Commun. 255:123-128(1999).  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 FAMILY.  
 DR EMBL: AF100931; AAD20218.1; -.  
 DR HSSP: P50163; 2AEL.  
 DR InterPro: IPR002198; -.  
 DR InterPro: IPR002347; -.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDHRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SQ SEQUENCE 276 AA; 30139 MW; 135F17111D30086D CRC64;

Query Match 3.8%; Score 8; DB 13; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATE 35  
 |||||  
 Db 93 AFKNDATE 100

## RESULT 5

O9RDF5 PRELIMINARY; PRT; 499 AA.

ID O9RDF5; PRELIMINARY; PRT; 499 AA.  
 AC O9RDF5;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE PROBABLE TRANSMEMBRANE TRANSPORT PROTEIN.  
 GN SC77.03C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:34:31 ; Search time 79.75 Seconds  
(without alignments)  
161.917 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 213  
Sequence: 1 MQLSLAPCLACLVAHAFAV.....KPPRRARGAKNAQALENNAY 213

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

riched: 412676 seqs, 60623988 residues

Word size: 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database :

A\_Geneseq\_0601:\*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:\*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	213	21	AAV96433
2	108	50.7	211	21	AAV96432
3	33	15.5	176	21	AAV96434
4	33	15.5	213	21	AAV96430
5	33	15.5	213	21	AAV96429
6	33	15.5	213	21	AAV96430
7	33	15.5	213	21	AAV96431
8	33	15.5	213	21	AAV96436
9	33	15.5	213	22	AAV97589
10	33	15.5	367	21	AAV97589
11	15	7.0	50	20	AAV12009

12	9	4.2	116	18	AAW27654
13	9	4.2	116	18	AAW44090
14	9	4.2	206	18	AAW09408
15	9	4.2	206	18	AAW58704
16	9	4.2	206	20	AAW95711
17	9	4.2	206	21	AAW10233
18	9	4.2	206	21	AAV75981
19	9	4.2	206	21	AAV76031
20	9	4.2	206	22	AAV55920
21	9	4.2	206	22	AAV55970
22	8	3.8	317	21	AAV58364
23	8	3.8	428	20	AAV52704
24	8	3.8	428	20	AAV32824
25	8	3.8	453	19	AAW54356
26	8	3.8	647	19	AAW48383
27	7	3.3	55	21	AAW40943
28	7	3.3	79	20	AAV73911
29	7	3.3	99	21	AAV68229
30	7	3.3	99	21	AAV52883
31	7	3.3	99	22	AAV56644
32	7	3.3	186	19	AAW65018
33	7	3.3	219	21	AAW42285
34	7	3.3	256	20	AAW88386
35	7	3.3	265	21	AAW42284
36	7	3.3	282	21	AAW42283
37	7	3.3	285	21	AAW29625
38	7	3.3	532	21	AAW14458
39	7	3.3	542	21	AAW14457
40	7	3.3	588	21	AAW14456
41	7	3.3	708	20	AAW88383
42	7	3.3	1099	22	AAW46768
43	7	3.3	1142	22	AAW46766
44	7	3.3	1142	22	AAW46766
45	7	3.3	1142	22	AAW46766
46	6	2.8	12	6	AAV50474
47	6	2.8	12	10	AAV92004
48	6	2.8	13	7	AAV60742
49	6	2.8	20	13	AAW27052
50	6	2.8	23	15	AAW57116

#### ALIGNMENTS

RESULT 1	
AAV96433	standard; Protein; 213 AA.
ID AAV96433	
AC AAV96433	
XX	
DT 12-SEP-2000	(first entry)
XX	
DE Rat TGF-beta binding protein (BEER).	
XX	
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; gene therapy; antisense therapy; fracture; bone mineralization.	
XX	
OS Rattus norvegicus.	
XX	
PN WO200032773-A1.	
XX	
PD 08-JUN-2000.	
XX	
PF 24-NOV-1999;	99WO-US27990.
XX	
PR 27-NOV-1998;	98US-0110283.
XX	
PA (DAWN-) DARWIN DISCOVERY LTD.	
XX	
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
XX Van Ness J, Winkler DG;	
XX WPI, 2000-412321/35.	
DR	

Secreted protein A  
Human secreted pro  
Human small CCN-1  
Human small CCN-1  
Human small CCN-1  
Homo sapiens fetal  
Human adult retina  
Murine skin cell p  
Murine skin cell p  
Skin cell protein,  
Skin cell protein,  
Lung cancer associ  
Human ELK-1 protei  
Human ELK-1. Homo  
Ubiquinolcytochrom  
Homo sapiens don-1  
Human OREF OREF707  
Human prostate tum  
Guinea pig beta-2-  
Guinea pig beta-2-  
Guinea pig beta-2m  
Forlythia dirigent  
Arabidopsis thalia  
Mammalian Zneul po  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Cat flea HMT perit  
Arabidopsis thalia  
Arabidopsis thalia  
Mammalian Zneul po  
FANCI2 protein fr  
FANCI2 protein fr  
FANCI2 protein fr  
FANCI2 protein fr  
Determinant site o  
Human prolactin C-  
Sequence from synt  
N-terminal prolact  
Immunostimulatory

DR N-PSDB: AAA29059.  
XX Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
PS Claim 6; Page 125-126; 162pp; English.  
XX  
XX This shows a rat transforming growth factor-beta (TGF-beta) binding  
CC protein designated rBEER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patient's own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
CC associated with low mineral content.  
XX  
XX Sequence 213 AA:  
S0  
Query Match 100.0%; Score 213; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 6,7e-215;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MGLSLACIACLLVHAFAVESGQWQAFKNDATETIIGLHRYPPPELENNOTNRAE 60  
DB 1 mqlslapclcllvhaafavesgqwkndatetlpglreypppeleennqtmrae 60  
OY 61 NGRPHPHRYDTDVSEYSCRELHYRFTDGPCKRSAPVTELVSQCGPARLLPNAIG 120  
DB 61 ngrpnhpdytdvseyscrelhyrftdgpckrsakpvtelevsgcgparllpnaig 120  
OY 121 RVKWRPNNGDFRCIPDRRAORVOLLCPGAAPRSRRVRLVASCKCKRLRFHNOSELK 180  
DB 121 rvkwrpnngdfrcipdryraqvllcp99aaprsrvrivaasckckrlrfhngselk 180  
OY 181 DFGFETARPOKRRPRRAGAKANQAELENAY 213  
DB 181 dfgfetarpokrrprrragakanqaeleay 213  
RESULT 2  
AA96432  
D AAY96432 standard; Protein: 211 AA.  
XX  
AC AAY96432:  
XX 12-SEP-2000 (first entry)  
XX  
DE Murine TGF-beta binding protein (BEER).  
XX  
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
XX  
OS Mus musculus.  
XX  
PN W0200032773-A1.  
XX  
PD 08-JUN-2000.  
XX  
PE 24-NOV-1999; 99WO-US27990.  
XX  
PR 27-NOV-1998; 98US-0110283.  
XX  
PA (DARW-) DARWIN DISCOVERY LTD.  
XX

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepert BW;  
PI Van Ness J, Winkler DG;  
XX  
XX WPI: 2000-412321/35.  
DR N-PSDB: AAA29058.  
XX  
XX Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
PS Claim 5; Page 124; 162pp; English.  
XX  
XX This shows a murine transforming growth factor-beta (TGF-beta)  
CC binding protein designated mBEER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patient's own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
CC associated with low mineral content.  
XX  
XX Sequence 211 AA:  
S0  
Query Match 50.7%; Score 108; DB 21; Length 211;  
Best Local Similarity 100.0%; Pred. No. 5.5e-105;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 90 TDGCRSAKPVTELVSGQCGPARLLPNAIGRVKWRPNNGDFRCIPDRRAORVOLLCP 149  
DB 88 tdgcrsakpvtelevsgcgparllpnaigrvkwrpnngdfrcipdryraqvllcp 147  
OY 150 GGAAPRSRRVRLVASCKCKRLRFHNOSELKDFGSETARPOKRRPR 197  
DB 148 ggaaprsrvrivaasckckrlrfhngselkdfigfetarpokrrpr 195  
RESULT 3  
AA96434  
D AAY96434 standard; Protein: 176 AA.  
XX  
AC AAY96434:  
XX 12-SEP-2000 (first entry)  
XX  
DE Bovine TGF-beta binding protein (BEER).  
XX  
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
XX  
OS Bos taurus.  
XX  
PN W0200032773-A1.  
XX  
PD 08-JUN-2000.  
XX  
PE 24-NOV-1999; 99WO-US27990.  
XX  
PR 27-NOV-1998; 98US-0110283.  
XX  
PA (DARW-) DARWIN DISCOVERY LTD.  
XX  
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepert BW;  
PI Van Ness J, Winkler DG;  
XX  
XX WPI: 2000-412321/35.  
DR



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:35:25 ; Search time 47.66 Seconds

(without alignments)  
100.571 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 213  
Sequence: 1 MOLSIAPCLACLVHAARVA.....KPRPRGAKANQALEENAY 213

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

Archived: 212252 seqs, 22503292 residues

Word size: 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database:

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5a\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5b\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/5a\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/5b\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	4.2	206	1	US-08-468-847B-2
2	9	4.2	206	1	US-08-468-847B-20
3	9	4.2	206	4	US-09-188-930-159
4	9	4.2	206	4	US-09-188-930-286
5	8	3.8	647	3	US-08-753-007A-32
6	8	3.8	647	4	US-09-398-496-32
7	7	3.3	99	2	US-08-484-905-61
8	7	3.3	99	3	US-08-481-985B-61
9	7	3.3	99	4	US-08-370-476-61
10	7	3.3	186	4	US-09-475-316A-13
11	7	3.3	861	1	US-08-764-100-24
12	6	2.8	24	1	US-08-220-272A-6
13	6	2.8	46	4	US-09-188-930-162
14	6	2.8	46	4	US-09-188-930-289
15	6	2.8	68	2	US-08-836-791-5
16	6	2.8	97	2	US-08-245-511-30
17	6	2.8	97	2	US-08-600-993A-30
18	6	2.8	125	3	US-08-985-526-25
19	6	2.8	128	1	US-08-666-798-2
20	6	2.8	128	1	US-08-892-692-2
21	6	2.8	128	1	US-09-096-071-2
22	6	2.8	199	3	US-08-737-248-7
23	6	2.8	202	4	US-09-342-284-10
24	6	2.8	207	4	US-08-652-877-13
25	6	2.8	207	4	US-08-476-515A-13
26	6	2.8	215	3	US-08-462-778-2
27	6	2.8	216	2	US-08-821-637-3

28	6	2.8	223	4	US-08-928-941D-22	Sequence 22, Appl
29	6	2.8	223	4	US-09-280-590A-22	Sequence 22, Appl
30	6	2.8	253	3	US-08-985-526-27	Sequence 27, Appl
31	6	2.8	256	2	US-08-727-688-33	Sequence 33, Appl
32	6	2.8	288	5	PCT-US92-00282-19	Sequence 19, Appl
33	6	2.8	330	4	US-09-188-930-144	Sequence 144, App
34	6	2.8	330	4	US-09-188-930-278	Sequence 278, App
35	6	2.8	343	2	US-08-933-750C-13	Sequence 13, Appl
36	6	2.8	343	4	US-09-234-613-13	Sequence 13, Appl
37	6	2.8	351	4	US-08-196-350-1	Sequence 1, Appl
38	6	2.8	351	4	US-09-245-041-11	Sequence 11, Appl
39	6	2.8	372	4	US-08-928-941D-16	Sequence 16, Appl
40	6	2.8	372	4	US-09-280-590A-16	Sequence 16, Appl
41	6	2.8	388	1	US-08-087-772A-2	Sequence 2, Appl
42	6	2.8	392	1	US-08-271-354-11	Sequence 11, Appl
43	6	2.8	392	2	US-08-565-861-11	Sequence 11, Appl
44	6	2.8	392	5	PCT-US94-07658-11	Sequence 11, Appl
45	6	2.8	395	3	US-08-981-825-6	Sequence 6, Appl
46	6	2.8	395	4	US-09-480-784-6	Sequence 6, Appl
47	6	2.8	400	1	US-07-916-901-6	Sequence 6, Appl
48	6	2.8	400	1	US-07-783-602C-1	Sequence 1, Appl
49	6	2.8	400	1	US-08-351-473B-4	Sequence 4, Appl
50	6	2.8	400	1	US-08-351-473B-5	Sequence 5, Appl

#### ALIGNMENTS

RESULT 1  
US-08-468-847B-2  
; Sequence 2, Application US/08468847B  
; Patent No. 5780263  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
; TITLE OF INVENTION: Human CCN-Like Growth Factor  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468, 847B  
; FILING DATE: 6 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33, 073  
; REFERENCE/DOCKET NUMBER: 325800-442  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-08-468-847B-2

Query Match 4.2%; Score 9; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 2

US-08-468-847B-20  
; Sequence 20, Application US/08468847B  
; Patent No. 5780263  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
; TITLE OF INVENTION: Human CCN-Like Growth Factor  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,847B  
; FILING DATE: 6 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-442  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
US-08-468-847B-20

Query Match 4.2%; Score 9; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 3

US-09-188-930-159  
; Sequence 159, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg

; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 159  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: mouse  
US-09-188-930-159

Query Match 4.2%; Score 9; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 4

US-09-188-930-286  
; Sequence 286, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 286  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-188-930-286

Query Match 4.2%; Score 9; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AFKNDATET 36  
Db 23 AFKNDATET 31

## RESULT 5

US-08-753-007A-32  
; Sequence 32, Application US/08753007A  
; Patent No. 6074841  
; GENERAL INFORMATION:  
; APPLICANT: Geating, David P.  
; APPLICANT: Busfield, Samantha J.  
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES  
; TITLE OF INVENTION: AND USES THEREFOR  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 9, 2001, 15:36:30 ; Search time 54.37 Seconds  
(without alignments) 246.583 Million cell updates/sec

Title: US-09-668-021-16

Sequence: 1 NDATTEILPELGEYPELPEL.....RPOTGKRLRRARGTKASRA 176

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

rchd: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	4.0	119	2	S14292
2	7	4.0	130	2	D83305
3	7	4.0	139	2	A82330
4	7	4.0	148	2	C83091
5	7	4.0	183	2	B37410
6	7	4.0	230	2	I37095
7	7	4.0	321	2	B82892
8	7	4.0	325	2	T35023
9	7	4.0	328	2	S72647
10	7	4.0	341	2	S71223
11	7	4.0	343	2	B86446
12	7	4.0	398	1	S24802
13	7	4.0	426	2	T15285
14	7	4.0	521	2	S54266
15	7	4.0	562	2	T05758
16	7	4.0	573	2	B70942
17	7	4.0	632	1	VGVSNY
18	7	4.0	864	1	JS0076
19	7	4.0	1168	1	MMAXIC
20	7	4.0	1308	2	T15280
21	7	4.0	1646	2	T40198
22	7	4.0	3006	2	T28625
23	7	4.0	6420	2	T10283
24	6	3.4	49	2	S28215
25	6	3.4	65	2	B25025
26	6	3.4	83	2	S41672
27	6	3.4	89	2	D75271
28	6	3.4	99	2	T17268
29	6	3.4	105	2	H72708

30	6	3.4	106	2	A82560	50S ribosomal prot
31	6	3.4	110	1	R5EC22	ribosomal protein
32	6	3.4	110	2	H85996	50S ribosomal subu
33	6	3.4	111	2	C41839	ribosomal protein
34	6	3.4	113	2	S66512	ribosomal protein
35	6	3.4	118	2	T15235	hypothetical prote
36	6	3.4	119	2	JQ2032	lambda 208 protein
37	6	3.4	124	1	R3EC12	ribosomal protein
38	6	3.4	124	2	JH0443	ribosomal protein
39	6	3.4	124	2	E85998	30S ribosomal subu
40	6	3.4	124	4	JH0807	ribosomal protein
41	6	3.4	124	4	JH0809	ribosomal protein
42	6	3.4	125	2	T10936	cellulase (EC 3.2.
43	6	3.4	129	2	T46388	hypothetical prote
44	6	3.4	130	2	E82647	hypothetical prote
45	6	3.4	134	2	D75534	ribosomal protein
46	6	3.4	136	2	T45360	hypothetical prote
47	6	3.4	137	2	PC4297	cadherin FIB1 - hu
48	6	3.4	141	2	B81268	50S ribosomal prot
49	6	3.4	142	2	H72600	hypothetical prote
50	6	3.4	145	2	G69011	hypothetical prote

## ALIGNMENTS

RESULT 1  
S14292  
transcription activator - *Emeritella nidulans*  
C:Species: *Emeritella nidulans*, *Aspergillus nidulans*  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 02-Jul-1998  
C:Accession: S14292  
R:Kulmburg, P.; Prange, T.; Mathieu, M.; Sequeval, D.; Scazzocchio, C.; Fellenbock, B.  
FEBS Lett. 280, 11-16, 1991  
A:Title: Correct intron splicing generates a new type of a putative zinc-binding doma  
A:Reference number: S14292; MUID:91184391  
A:Accession: S14292  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-119 <KUL>  
C:Superfamily: GAL4 zinc binuclear cluster homology  
F:7-54/Domain: GAL4 zinc binuclear cluster homology <GAL4>

Query Match 4.0%; Score 7; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 GAAPRAR 126  
|||||||  
Db 63 GAAPRAR 69

RESULT 2  
D83305  
hypothetical protein PA2722 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83305  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337  
A:Accession: D83305  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-130 <STO>  
A:Cross-references: GB:AE004700; GB:AE004091; MTD:99948792; PIDN:AMG06110.1; GSPDB:GN  
C:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2722

```
Query Match 4.0%; Score 7; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 KVRIVAS 133
    |||||
Db 12 KVRIVAS 18

RESULT 3
A82330
Conserved hypothetical protein VC0373 [Imported] - Vibrio cholerae (strain N16961 serogr
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82330
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Charlton, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F
I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: A82330
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-139 <HEI>
A:Cross-references: GB:AE004126; GB:AE003852; NID:99654802; PIDN:AAF93546.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0373
A:Map position: 1
C:Superfamily: hypothetical protein MJ1081

Query Match 4.0%; Score 7; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 LRPRARG 170
    |||||
Db 9 LRPRARG 15

RESULT 4
C83091
hypothetical protein PA4441 [Imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83091
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bu
dman, S.; Olson, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: A82950; MUID:20437337
A:Accession: C83091
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <STO>
A:Cross-references: GB:AE004858; GB:AE004091; NID:99950668; PIDN:AG07829.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4441

Query Match 4.0%; Score 7; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 ARILPNA 87
    |||||
Db 26 ARILPNA 32
```

```
RESULT 5
B37410
H-2 class II histocompatibility antigen A-2 beta chain - spiny mouse (Mus saxicola) (
C:Species: Mus saxicola
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
C:Accession: B37410
R:Cam, P.; Jovin-Marche, E.; Leguern, C.; Marche, P.N.
Eur. J. Immunol. 20, 1337-1343, 1990
A>Title: Structure of class II genes in wild mouse Mus saxicola: functional and evolu
A:Reference number: A37410; MUID:90316177
A:Accession: B37410
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-183 <CAM>
A:Cross-references: GB:M30158
C:Superfamily: class II histocompatibility antigen: immunoglobulin homology
F:105-170/Domain: immunoglobulin homology <IMM>

Query Match 4.0%; Score 7; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EIIPELG 11
    |||||
Db 157 EIIPELG 163

RESULT 6
I37095
gene 2.19 protein - human
C:Species: Homo sapiens (man)
C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C:Accession: I37095
R:Bione, S.; Tamarini, F.; Maestri, E.; Tribioli, C.; Poustka, A.; Torri, G.; Rivet
Proc. Natl. Acad. Sci. U.S.A. 90, 10977-10981, 1993
A>Title: Transcriptional organization of a 450-kb region of the human X chromosome in
A:Reference number: I37095; MUID:94068527
A:Accession: I37095
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-230 <RES>
A:Cross-references: EMBL:X87193; NID:9854081; PIDN:CAA60645.1; PID:9854082
C:Genetics:
A:Gene: 2.19
C:Superfamily: human gene 2.19 protein

Query Match 4.0%; Score 7; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 AAPRARK 127
    |||||
Db 50 AAPRARK 56

RESULT 7
B82892
Conserved hypothetical U0417 [Imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82892
R:Glass, J.I.; Letikowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A>Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: B82892
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <GLA>
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:58 ; Search time 31.03 Seconds  
(without alignments)  
194.295 Million cell updates/sec

Title: US-09-668-021-16

Perfect score: 176  
Sequence: 1 NDATETIPELGEYEPPELP...RPQTGRKLPRPRAGTKASRA 176

Scoring table:  
Gapop 60.0 , Gapept 60.0

Archived: 93435 seqs, 34255486 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	4.0	110	1	RL22_LEPIN
2	7	4.0	230	1	219_HUMAN
3	7	4.0	632	1	VGEG_SYNY
4	7	4.0	821	1	ALCR_EMEI
5	7	4.0	1168	1	MYSC_ACACA
6	6	3.4	49	1	TX25_PHONI
7	6	3.4	65	1	MALX_KLEPN
8	6	3.4	97	1	WE7_HPVA4
9	6	3.4	105	1	RL31_AERPE
10	6	3.4	110	1	RL22_ECOTI
11	6	3.4	111	1	RL22_ACHIA
12	6	3.4	112	1	RL22_SPICI
13	6	3.4	113	1	RL22_THERH
14	6	3.4	115	1	TRAF_HUMAN
15	6	3.4	115	1	TRAF_MOUSE
16	6	3.4	119	1	REV_HYLLM
17	6	3.4	119	1	Y13K_NPYOP
18	6	3.4	123	1	RS12_ECOTI
19	6	3.4	131	1	RL22_PHYSI
20	6	3.4	148	1	RR22_MAIZE
21	6	3.4	149	1	RR22_ORYSA
22	6	3.4	150	1	R19E_PYRHO
23	6	3.4	167	1	B3AR_MERIN
24	6	3.4	182	1	YCY0_YEAST
25	6	3.4	184	1	MPL_MPLY
26	6	3.4	193	1	HS72_CANAL
27	6	3.4	196	1	RETB_CHICK
28	6	3.4	210	1	TRPF_KLUDA
29	6	3.4	211	1	DEOC_BACSU
30	6	3.4	216	1	RGFH_HUMAN
31	6	3.4	216	1	RGFH_MOUSE
32	6	3.4	227	1	PRL_HUMAN
33	6	3.4	227	1	PRL_MOUSE

34	6	3.4	229	1	YT67_CABEI	011083 caenorhabdi
35	6	3.4	230	1	YCAP_ECOTI	P75839 escherichia
36	6	3.4	247	1	PS72_XENLA	O9P941 xenopus lae
37	6	3.4	248	1	PS71_XENLA	O9P946 xenopus lae
38	6	3.4	248	1	PSA7_HUMAN	O14818 homo sapien
39	6	3.4	248	1	PSA7_MOUSE	O92200 mus musculu
40	6	3.4	249	1	PSA7_CHICK	O13268 gallus galli
41	6	3.4	250	1	YL27_YEAST	O07821 saccharomyc
42	6	3.4	252	1	MYBD_MAIZE	P23592 zea mays (m
43	6	3.4	253	1	TPIS_BACST	P00943 bacillus st
44	6	3.4	254	1	PSA7_RAT	P48004 rattus norv
45	6	3.4	268	1	ISPE_AQUAE	O67060 aquilex aeo
46	6	3.4	269	1	ESLI_MYCPN	P75333 mycoplasma
47	6	3.4	274	1	PANB_AERPE	O9Y697 aeropyrum p
48	6	3.4	280	1	HESI_HUMAN	O14469 homo sapien
49	6	3.4	281	1	HESI_RAT	O04666 rattus norv
50	6	3.4	282	1	HESI_MOUSE	P35428 mus musculu

## ALIGNMENTS

```

RESULT 1
RL22_LEPIN
ID RL22_LEPIN STANDARD: PRT: 110 AA.
AC Q9XD31
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L22.
GN RPLV.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_Taxid=173;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SERVAR LAI;
RX MEDLINE=20088835; PubMed=10620683;
RA Zuercher R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;
RT Characterization of the Leptospira interrogans S10-spc-alpha
operon ".
RL FEMS Microbiol. Lett. 182:303-308(2000).
CC -I- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S RNA; ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., LA, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF115283; AAD40588.1; -.
DR InterPro: IPR001063; -.
DR Pfam: PF00237; Ribosomal_L22; 1.
DR PROSITE: PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; rRNA-binding.
SQ
SEQUENCE 110 AA; 12566 MW; B9181B85E9756CAA CRC64;

```

Query Match 4.0%; Score 7; DB 1; Length 110;  
Best local Similarity 100.0%; Pred. No. 3.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 126 RKVRLVA 132  
DB 15 RKVRLVA 21

```

RESULT 2
ID 219_HUMAN STANDARD; PRT; 230 AA.
AC P88173;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 2-19 PROTEIN PRECURSOR.
GN 2-19 OR 2.19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA MEDLINE=94068527; PubMed=8248200;
RA Bione S., Tamameli F., Maestri E., Tribioli C., Poustka A.,
RA Torri G., Rivella S., Toniolo D.;
RA "Transcriptional organization of a 450-kb region of the human X
RA chromosome in Xq28.";
RA Proc. Natl. Acad. Sci. U.S.A. 90:10977-10981(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Zollo M., Mazzarella R., Bione S., Toniolo D., Schlessinger D.,
RA D'Urso M., Chen E.Y.;
RA Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA PubMed=873315;
RA Chen E.Y., Zollo M., Mazzarella R.A., Ciccodicola A., Chen C.-N.,
RA Zuo L., Heiner C., Burrough F.W., Ripetto M., Schlessinger D.,
RA D'Urso M.;
RA "Long-range sequence analysis in Xq28: thirteen known and six
RA candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and
RA Hum. Mol. Genet. 5:659-668(1996).
RN [4]
RP TISSUE SPECIFICITY: IN SIMILAR AMOUNTS IN TESTIS, PANCREAS,
RN ADRENAL, PLACENTA, BRAIN, FETAL BRAIN, LIVER, KIDNEY, SKELETAL
RN MUSCLE, HEART.
CC -1 SIMILARITY: BELONGS TO THE 2-19 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X55448; CA39090.1; -;
EMBL; X87193; CA60645.1; -;
DR EMBL; L44140; AAA92652.1; -;
KW Signal.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 230 2-19 PROTEIN.
SQ SEQUENCE 230 AA; 25069 MW; FE3934D91F98CAD CRC64;

Query Match 4.0%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 AAPPARK 127
DB 50 AAPPARK 56

RESULT 3
VGLG_SYNV STANDARD; PRT; 632 AA.
ID VGLG_SYNV
AC P27277;
DT 01-AUG-1992 (Rel. 23, Created)

```

```

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE SPIKE GLYCOPROTEIN PRECURSOR.
GN G.
OS Sonchus yellow net virus (SYNV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Nucleorhabdovirus.
CX NCBI_TaxID=11307;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC PV-263;
RX MEDLINE=92024089; PubMed=1926779;
RA Goldberg K.B., Modrell B., Hillman B.I., Heaton L.A., Choi T.J.,
RA Jackson A.O.;
RA "Structure of the glycoprotein gene of sonchus yellow net virus, a
RA plant rhabdovirus.";
RL Virology 185:32-38(1991).
CC -1 FUNCTION: THIS PROTEIN FORMS SPIKES ON THE SURFACE OF THE VIRION.
CC IT IS RESPONSIBLE BOTH FOR THE BINDING OF THE VIRUS TO SUSCEPTIBLE
CC HOST CELLS AND FOR INDUCING THE UPTAKE OF THE VIRUS BY THE CELL.
CC THE INTERACTION BETWEEN THE INTERNAL COMPONENTS OF THE VIRION
CC AND THE PORTION OF THE GLYCOPROTEIN EXPOSED ON THE CYTOPLASMIC
CC FACE OF THE PLASMA MEMBRANE PROBABLY DIRECTS ENVELOPMENT AND
CC VIRUS BUDDING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32603; AAA50384.1; -;
DR EMBL; M73626; AAA47898.1; -;
DR PIR; A40776; VGVNSY.
KW Transmembrane; Envelope protein; Glycoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 632 SPIKE GLYCOPROTEIN.
FT TRANSMEM 560 578 POTENTIAL.
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 512 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 632 AA; 71093 MW; 05541E78BEA07927 CRC64;

Query Match 4.0%; Score 7; DB 1; Length 632;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 KVLVANS 133
DB 579 KVLVANS 585

RESULT 4
ALCR_EMENTI STANDARD; PRT; 821 AA.
ID ALCR_EMENTI
AC P21228;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE REGULATORY PROTEIN ALCR.
GN ALCR.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
CX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89211976; PubMed=3072264;

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:47:21; Search time 89.8 Seconds  
(without alignments)  
259.306 Million cell updates/sec

Title: US-09-668-021-16

Perfect score: 176

Sequence: 1 NDATETIPELGEYEPPELPEL.....RPQGRKLPRRANGTKASRA 176

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Search: 425026 seqs, 132305027 residues

Word size: 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 50 summaries

Database:

1: SPREMBL\_16:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_ricent:\*  
13: sp\_unclassified:\*  
14: sp\_vertebrate:\*  
15: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	4.0	45	4	Q9HC30
2	7	4.0	89	10	Q9PH55
3	7	4.0	106	14	Q9IGW4
4	7	4.0	109	5	Q9GTJ4
5	7	4.0	130	2	Q9I0C0
6	7	4.0	139	2	Q9KUY5
7	7	4.0	148	2	Q9HXY4
8	7	4.0	154	9	Q21999
9	7	4.0	183	7	Q31243
10	7	4.0	220	2	Q9RKY8
11	7	4.0	239	14	Q9IBJ1
12	7	4.0	282	10	Q9XF19
13	7	4.0	321	2	Q9PG72
14	7	4.0	325	2	Q9XAM5
15	7	4.0	341	10	Q38908
16	7	4.0	343	10	Q9FVR2
17	7	4.0	362	2	Q9RUX3
18	7	4.0	398	1	Q00388
19	7	4.0	425	5	Q19348

20	7	4.0	426	5	Q01969	Q01969 caenorhabd1
21	7	4.0	438	2	Q9EZJ8	Q9EZJ8 thermus agu
22	7	4.0	442	11	Q35752	Q35752 rattus norv
23	7	4.0	468	4	Q60276	Q60276 homo sapien
24	7	4.0	493	11	Q9UMC2	Q9UMC2 mus musculu
25	7	4.0	521	14	Q66044	Q66044 caprine her
26	7	4.0	523	10	Q9M422	Q9M422 hordue vul
27	7	4.0	562	10	Q81837	Q81837 arabidopsis
28	7	4.0	573	2	Q53473	Q53473 mycobacteri
29	7	4.0	604	5	Q44003	Q44003 toxoplasma
30	7	4.0	620	5	Q2NEC8	Q2NEC8 leishmania
31	7	4.0	638	4	Q9H7T7	Q9H7T7 homo sapien
32	7	4.0	643	14	Q85050	Q85050 pseudorabie
33	7	4.0	825	4	Q9H3H2	Q9H3H2 homo sapien
34	7	4.0	1186	5	Q61080	Q61080 acanthamoeb
35	7	4.0	1308	5	Q01924	Q01924 caenorhabd1
36	7	4.0	1325	2	Q9WK63	Q9WK63 acetobacter
37	7	4.0	1646	3	Q94649	Q94649 schizosacch
38	7	4.0	3006	5	Q26032	Q26032 plasmodium
39	7	4.0	6420	2	Q95814	Q95814 streptomyce
40	6	3.4	22	2	Q85607	Q85607 streptomyce
41	6	3.4	45	6	Q9GK71	Q9GK71 bos taurus
42	6	3.4	60	14	Q69367	Q69367 cercopithec
43	6	3.4	62	10	Q9LDG4	Q9LDG4 oryza sativ
44	6	3.4	64	1	Q9HJ81	Q9HJ81 thermoplasm
45	6	3.4	68	5	Q9NMG9	Q9NMG9 leishmania
46	6	3.4	76	14	Q68828	Q68828 human cytom
47	6	3.4	89	2	Q9RMO0	Q9RMO0 delinococcus
48	6	3.4	91	14	Q72747	Q72747 human immun
49	6	3.4	91	14	Q72751	Q72751 human immun
50	6	3.4	91	14	Q72753	Q72753 human immun

## ALIGNMENTS

RESULT 1  
ID Q9HC30 PRELIMINARY; PRT; 45 AA.  
AC Q9HC30;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CAMP SPECIFIC PHOSPHODIESTERASE VARIANT PDE4A-10 (FRAGMENT).  
GN PDE4A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sullivan M., Rena G., Begg F., Olsen A.S., Houslay M.D.;  
RT "PDE4A-10, a novel human 5' splice variant of the CAMP specific  
RT phosphodiesterase PDE4 gene.";  
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AF178570; AAC13806.1; -  
FT NON-TER  
FT SEQUENCE 45 AA; 4887 MW; A9A65DE550D1283D CRC64;

Query Match 4.0%; Score 7; DB 4; Length 45;  
Best Local Similarity 100.0%; Pred. No. 6.3; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 120 GAAPRAR 126  
IIIIII  
Db 4 GAAPRAR 10

RESULT 2  
ID Q9FH55 PRELIMINARY; PRT; 89 AA.  
AC Q9FH55;

DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K1120.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty p1 and TAC  
 RT clones.";  
 RL DNA Res. 7:31-63(2000).  
 DR EMBL:AB022211; BAB10713.1; -  
 SQ SEQUENCE 89 AA; 10345 MW; DF1234C836FAEDA CRC64;

Query Match 4.0%; Score 7; DB 10; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 RLPLNAT 88  
 Db 18 RLPLNAT 24

RESULT 3  
 ID Q9IGM4 PRELIMINARY; PRT; 106 AA.  
 AC Q9IGM4;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE P12.  
 GN K19.  
 OS Hypoviridae; dsDNA viruses, no RNA stage; Polydnaviridae; Ichnovirus.  
 OC NCBI\_TaxID=56779;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Volkoff N.A., Roehrer J., Cernutti P., Ohresser M.C.,  
 RA d'Aubenton-Carafa Y., Devauchelle G., Duonor-Cernutti M.;  
 RT "Persistent expression of a newly characterized Hypoviridae didymator  
 RT POLYNAVIRUS gene in long-term infected Lepidopteran cell lines.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL:AF241775; AAF91314.1; -  
 SQ SEQUENCE 106 AA; 11821 MW; D810A541DA90A29C CRC64;

Query Match 4.0%; Score 7; DB 14; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 PEPLPEL 20  
 Db 18 PEPLPEL 24

RESULT 4  
 ID Q9GTJ4 PRELIMINARY; PRT; 109 AA.  
 AC Q9GTJ4;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE ZINC METALLOPEPTIDASE 3 MEPS (FRAGMENT).  
 OS Ancylostoma caninum (dog hookworm).

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
 OC NCBI\_TaxID=29170;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-WHOLE WORMS;  
 RA Jones B.F., Hotez P.J.;  
 RT "Cloning and characterization of a zinc-metalloprotease secreted by  
 RT the invasive stages of Ancylostoma caninum";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF273706; AAG29106.1; -  
 FT NON-TER 1 1  
 SQ SEQUENCE 109 AA; 12300 MW; 7C3C54FD724CAD7B CRC64;

Query Match 4.0%; Score 7; DB 5; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 100 DFCICPD 106  
 Db 86 DFCICPD 92

RESULT 5  
 ID Q910C0 PRELIMINARY; PRT; 130 AA.  
 AC Q910C0;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA2722.  
 GN PA2722.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OC NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Madman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Loy S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004700; AAG06110.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 130 AA; 14641 MW; ECEBDCD7F247836E CRC64;

Query Match 4.0%; Score 7; DB 2; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 KVLVNS 133  
 Db 12 KVLVNS 18

RESULT 6  
 ID Q9KUY5 PRELIMINARY; PRT; 139 AA.  
 AC Q9KUY5;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE HYPOTHETICAL PROTEIN VC0373.  
 GN VC0373.



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:34:32 ; Search time 79.75 Seconds

Title:	US-09-668-021-16
Perfect score:	176
Sequence:	1 NDATETIPELGEYEPPLPEL.....RPQTGRKLRPRARGTKASRA 176

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 50 summaries

Database : A\_Geneseq\_0601:\*

1:	/S1D8_gcgdata/genseq/genseq/A11960.DAT*
2:	/S1D8_gcgdata/genseq/genseq/A11961.DAT*
3:	/S1D8_gcgdata/genseq/genseq/A11962.DAT*
4:	/S1D8_gcgdata/genseq/genseq/A11963.DAT*
5:	/S1D8_gcgdata/genseq/genseq/A11964.DAT*
6:	/S1D8_gcgdata/genseq/genseq/A11965.DAT*
7:	/S1D8_gcgdata/genseq/genseq/A11966.DAT*
8:	/S1D8_gcgdata/genseq/genseq/A11967.DAT*
9:	/S1D8_gcgdata/genseq/genseq/A11968.DAT*
10:	/S1D8_gcgdata/genseq/genseq/A11969.DAT*
11:	/S1D8_gcgdata/genseq/genseq/A11970.DAT*
12:	/S1D8_gcgdata/genseq/genseq/A11971.DAT*
13:	/S1D8_gcgdata/genseq/genseq/A11972.DAT*
14:	/S1D8_gcgdata/genseq/genseq/A11973.DAT*
15:	/S1D8_gcgdata/genseq/genseq/A11974.DAT*
16:	/S1D8_gcgdata/genseq/genseq/A11975.DAT*
17:	/S1D8_gcgdata/genseq/genseq/A11976.DAT*
18:	/S1D8_gcgdata/genseq/genseq/A11977.DAT*
19:	/S1D8_gcgdata/genseq/genseq/A11978.DAT*
20:	/S1D8_gcgdata/genseq/genseq/A11979.DAT*
21:	/S1D8_gcgdata/genseq/genseq/A12000.DAT*
22:	/S1D8_gcgdata/genseq/genseq/A12001.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	176	100.0	176	21	AAV96434	Bovine TGF-beta b1
2	115	65.3	213	21	AAV96431	Vervet TGF-beta b1
3	76	43.2	213	21	AAV826106	Human DNF/Cerberus
4	76	43.2	213	21	AAV96429	Human TGF-beta b1n
5	76	43.2	213	21	AAV96430	Human TGF-beta b1n
6	76	43.2	213	21	AAV96436	Human TGF-beta b1n
7	76	43.2	213	22	AAV97589	Human secreted prec
8	76	43.2	367	21	AAV826105	Human DNF/Cerberus
9	33	18.8	213	21	AAV96433	Rat TGF-beta b1n
10	32	18.2	211	21	AAV96432	Murine TGF-beta b1
11	16	9.1	50	20	AAV12009	Human 5' EST secret

1	Arabidopsis thaliana	1	219	21	AA542285
2	Human cytokine fam	7	4.0	13	AA57845
3	Arabidopsis thaliana	7	4.0	230	AA57845
4	Arabidopsis thaliana	7	4.0	264	AA515729
5	Arabidopsis thaliana	7	4.0	265	AA542782
6	Arabidopsis thaliana	7	4.0	273	AA515728
7	Arabidopsis thaliana	7	4.0	282	AA542283
8	Arabidopsis thaliana	7	4.0	311	AA562150
9	Arabidopsis thaliana	7	4.0	343	AA515727
10	Arabidopsis thaliana	7	4.0	12	AA50474
11	Arabidopsis thaliana	7	4.0	12	AA50474
12	Arabidopsis thaliana	7	4.0	12	AA50474
13	Arabidopsis thaliana	7	4.0	12	AA50474
14	Arabidopsis thaliana	7	4.0	12	AA50474
15	Arabidopsis thaliana	7	4.0	12	AA50474
16	Arabidopsis thaliana	7	4.0	12	AA50474
17	Arabidopsis thaliana	7	4.0	12	AA50474
18	Arabidopsis thaliana	7	4.0	12	AA50474
19	Arabidopsis thaliana	7	4.0	12	AA50474
20	Arabidopsis thaliana	7	4.0	12	AA50474
21	Arabidopsis thaliana	7	4.0	12	AA50474
22	Arabidopsis thaliana	7	4.0	12	AA50474
23	Arabidopsis thaliana	7	4.0	12	AA50474
24	Arabidopsis thaliana	7	4.0	12	AA50474
25	Arabidopsis thaliana	7	4.0	12	AA50474
26	Arabidopsis thaliana	7	4.0	12	AA50474
27	Arabidopsis thaliana	7	4.0	12	AA50474
28	Arabidopsis thaliana	7	4.0	12	AA50474
29	Arabidopsis thaliana	7	4.0	12	AA50474
30	Arabidopsis thaliana	7	4.0	12	AA50474
31	Arabidopsis thaliana	7	4.0	12	AA50474
32	Arabidopsis thaliana	7	4.0	12	AA50474
33	Arabidopsis thaliana	7	4.0	12	AA50474
34	Arabidopsis thaliana	7	4.0	12	AA50474
35	Arabidopsis thaliana	7	4.0	12	AA50474
36	Arabidopsis thaliana	7	4.0	12	AA50474
37	Arabidopsis thaliana	7	4.0	12	AA50474
38	Arabidopsis thaliana	7	4.0	12	AA50474
39	Arabidopsis thaliana	7	4.0	12	AA50474
40	Arabidopsis thaliana	7	4.0	12	AA50474
41	Arabidopsis thaliana	7	4.0	12	AA50474
42	Arabidopsis thaliana	7	4.0	12	AA50474
43	Arabidopsis thaliana	7	4.0	12	AA50474
44	Arabidopsis thaliana	7	4.0	12	AA50474
45	Arabidopsis thaliana	7	4.0	12	AA50474
46	Arabidopsis thaliana	7	4.0	12	AA50474
47	Arabidopsis thaliana	7	4.0	12	AA50474
48	Arabidopsis thaliana	7	4.0	12	AA50474
49	Arabidopsis thaliana	7	4.0	12	AA50474
50	Arabidopsis thaliana	7	4.0	12	AA50474

## ALIGNMENTS

## RESULT 1

ID	AA	standard; 176 AA
1	AA	176

AC AAY96434

DT 12-SEP-2000 (first entry)  
XX

DE Bovine TGF-beta binding protein (BEER)  
XY

KM osteopalic; transforming growth factor-beta; TGF-beta; binding protein.  
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

05 Bos taurus.

PN WO200032773-A1

PD 08-JUN-2000.  
XY

XX 24-NOV-1999; 99WU-0527990.

FR 27-NOV-1998, 3603-0110263,  
XX XX

( ) DIRECTOR  
DIRECTOR OF THE FBI  
WASHINGTON, D.C.  
XX

XX

PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.

DR WPI; 2000-412321/35.

DR N-PSDB: AAA29060.  
 XX Nucleic acids (1) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 PS Claim 7; Page 127; 162pp; English.  
 XX This shows a bovine transforming growth factor-beta (TGF-beta)  
 CC binding protein designated bBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BBR expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BBR by  
 CC expressing inactive proteins or to supplement the patient's own production  
 CC of BBR polypeptides. The nucleic acids may be used for recombinant  
 CC production of BBR, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BBR may be used to raise  
 CC antibodies and for identification of BBR modulators. BBR antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 XX Sequence 176 AA:

Query Match 100.0%; Score 176; DB 21; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-174;  
 Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDAETIRPELGEPYRPELNNKTMNRAENGARPHHREMTDASEYSCRELHFRVYTD 60  
 DB 1 ndaetirpeigeypepelpelnkumraenggrprhphletkdaseyscrelhtlrvytd 60  
 QY 61 GPCRSAPVTELVCSCGCPARLLPNAIGRKWMPSPGDFRCIPDRYRAQRYOLLCPG 120  
 DB 61 gpcrsakpvtelvcsgcgparllpnaigrkwmrpsgdfrcicldryraqrvllcpgg 120  
 QY 121 AAPRAKRVLVASCCKRLRFRHNOSELKDFGEPAAROTGKRLPRRGRKASKA 176  
 DB 121 aaprakrvlvascckrlrfrhngselkdfgeaarpotgkrlprgrgkaskas 176

RESULT 2  
 AAY96431  
 ID AAY96431 standard; Protein: 213 AA.

AC AAY96431:  
 XX 12-SEP-2000 (first entry)  
 DE Vervet TGF-beta binding protein (bBER).  
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 KW bBER; gene therapy; antisense therapy; fracture; bone mineralization.  
 XX Cercopithecus pygerythrus.  
 OS WO200032773-A1.  
 PN 08-JUN-2000.  
 PD 24-NOV-1999; 99MO-US27990.  
 PF 27-NOV-1998; 98US-0110283.  
 PR (DARW-) DARWIN DISCOVERY LTD.  
 PA Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW,  
 PI Van Ness J, Winkler DG;  
 XX WPI: 2000-412321/35.

DR N-PSDB: AAA29057.  
 XX Nucleic acids (1) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 PS Claim 4; Page 122-123; 162pp; English.  
 XX This shows a vervet transforming growth factor-beta (TGF-beta)  
 CC binding protein designated vBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BBR expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BBR by  
 CC expressing inactive proteins or to supplement the patient's own production  
 CC of BBR polypeptides. The nucleic acids may be used for recombinant  
 CC production of BBR, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BBR may be used to raise  
 CC antibodies and for identification of BBR modulators. BBR antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 XX Sequence 213 AA:

Query Match 65.3%; Score 115; DB 21; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 3e-111;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 SEYSCRELHFTYVVDGCRSAKPYTELVCSCGCPARLLPNAIGRKWMPSPGDFRCI 104  
 DB 76 seyrcelhtfryvvdgcrsakpytelvcsgcgparllpnaigrkwmrpsgdfrcld 135  
 QY 105 PDYRAQRYOLLCPGGAAPRAKRVLVASCCKRLRFRHNOSELKDFGEPAARQ 159  
 DB 136 pdyraqrvllcpggaaaprakrvlvascckrlrfrhngselkdfgeaarpq 190

RESULT 3  
 AAB26106  
 ID AAB26106 standard; Protein: 213 AA.

AC AAB26106:  
 XX 15-JAN-2001 (first entry)  
 DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).  
 XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;  
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;  
 XX gene therapy.  
 OS Homo sapiens.  
 PN WO200055193-A2.  
 PD 21-SEP-2000.  
 PF 02-MAR-2000; 2000MO-US05537.  
 PR 12-MAR-1999; 99US-0124118.  
 PA (REGG-) REGENERON PHARM INC.  
 PI Economides AN;  
 XX WPI: 2000-638179/61.  
 DR N-PSDB: AAA94051.  
 PT Novel isolated, human DNA/Cerberus related protein 6 which include  
 natural homologue, and polypeptides comprising DCR6 domain and nucleic

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 9, 2001, 15:35:25 ; Search time 47.66 Seconds  
(without alignments)  
83.101 Million cell updates/sec

Title: US-09-668-021-16

Perfect score: 176  
Sequence: 1 NDAETIPELGEYEPPELPEL.....RPTGRKLRPRARGTAKSRA 176

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

rchcd: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

- Issued\_Patents\_AA.\*
- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*
  - 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*
  - 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep.\*
  - 6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	4.0	855	4	US-09-813-819-2
2	6	3.4	32	2	US-08-078-311-25
3	6	3.4	32	2	US-08-460-402-25
4	6	3.4	45	2	US-08-078-311-10
5	6	3.4	45	2	US-08-460-402-10
6	6	3.4	68	2	US-08-836-791-5
7	6	3.4	125	3	US-08-985-526-25
8	6	3.4	128	1	US-08-666-798-2
9	6	3.4	128	1	US-08-892-692-2
10	6	3.4	128	2	US-09-096-071-2
11	6	3.4	184	2	US-08-078-311-4
12	6	3.4	184	2	US-08-460-402-4
13	6	3.4	199	3	US-08-737-248-7
14	6	3.4	202	4	US-09-342-084-10
15	6	3.4	206	1	US-08-468-847B-2
16	6	3.4	206	1	US-08-468-847B-20
17	6	3.4	206	4	US-09-188-930-159
18	6	3.4	206	4	US-09-188-930-286
19	6	3.4	216	2	US-08-821-637-3
20	6	3.4	223	4	US-08-928-941D-22
21	6	3.4	223	4	US-09-280-590A-22
22	6	3.4	243	2	US-08-319-376-2
23	6	3.4	248	2	US-08-701-935-3
24	6	3.4	248	2	US-09-134-931-3
25	6	3.4	251	2	US-08-704-931-4
26	6	3.4	253	3	US-08-985-526-27
27	6	3.4	254	2	US-08-701-935-6

28	6	3.4	254	3	US-09-134-591-6	Sequence 6, Appl1
29	6	3.4	256	2	US-08-727-688-33	Sequence 33, Appl1
30	6	3.4	284	2	US-08-078-311-14	Sequence 14, Appl1
31	6	3.4	284	2	US-08-078-311-24	Sequence 24, Appl1
32	6	3.4	284	2	US-08-460-402-14	Sequence 14, Appl1
33	6	3.4	284	2	US-08-460-402-24	Sequence 24, Appl1
34	6	3.4	305	2	US-09-190-821-2	Sequence 2, Appl1
35	6	3.4	330	4	US-09-188-930-144	Sequence 144, App
36	6	3.4	330	4	US-09-188-930-278	Sequence 278, App
37	6	3.4	343	2	US-08-933-750C-13	Sequence 13, Appl1
38	6	3.4	343	4	US-09-234-613-13	Sequence 13, Appl1
39	6	3.4	349	4	US-09-343-011B-1	Sequence 1, Appl1
40	6	3.4	351	1	US-08-196-350-1	Sequence 1, Appl1
41	6	3.4	372	4	US-08-928-941D-16	Sequence 16, Appl1
42	6	3.4	382	2	US-09-280-590A-16	Sequence 16, Appl1
43	6	3.4	382	2	US-08-078-311-3	Sequence 3, Appl1
44	6	3.4	382	2	US-08-460-402-3	Sequence 3, Appl1
45	6	3.4	386	3	US-08-972-902-3	Sequence 3, Appl1
46	6	3.4	388	1	US-08-087-772A-2	Sequence 2, Appl1
47	6	3.4	392	1	US-08-271-354-11	Sequence 11, Appl1
48	6	3.4	392	2	US-08-565-861-11	Sequence 11, Appl1
49	6	3.4	392	5	PCT-US84-07658-11	Sequence 11, Appl1
50	6	3.4	395	3	US-08-981-825-6	Sequence 6, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-813-819-2
; Sequence 2, Application US/09813819
; Patent No. 6294368
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USRS THEREOF
; FILE REFERENCE: C1001177
; CURRENT APPLICATION NUMBER: US/09/813,819
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Human
; US-09-813-819-2

Query Match          4.0%; Score 7; DB 4; Length 855;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ARGTKAS 174
      |||||
Db 772 ARGTKAS 778

RESULT 2
US-08-078-311-25
; Sequence 25, Application US/08078311
; Patent No. 5925750
; GENERAL INFORMATION:
; APPLICANT: Charon, Martine
; APPLICANT: Gisselbrecht, Sylvie
; APPLICANT: Penciolelli, Jean-Francois
; APPLICANT: Souyri, Michele
; APPLICANT: Tambourin, Pierre
; APPLICANT: Varlet, Paule
; APPLICANT: Vigon, Isabelle
; APPLICANT: Wendling, Francoise
; TITLE OF INVENTION: Polypeptide of a Growth Factor Receptor
; TITLE OF INVENTION: Family, Application in the Diagnosis and Treatment of
```

TITLE OF INVENTION: Myeloproliferative Disease  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5925750west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/078,311  
FILING DATE: 18-JUN-1993  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR90/00762  
FILING DATE: 19-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 8076.84USMO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: Extracellular domain of v-mp1  
US-08-078-311-25  
Query Match 3.4%; Score 6; DB 2; Length 32;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 164 LRPRAR 169  
DB 2 LRPRAR 7  
RESULT 3  
US-08-460-402-25  
Sequence 25, Application US/08460402  
Patent No. 5989833  
GENERAL INFORMATION:  
APPLICANT: Charon, Martine  
APPLICANT: Gisselbrecht, Silvie  
APPLICANT: Penciolelli, Jean-Francis  
APPLICANT: Souvri, Michele  
APPLICANT: Tambourin, Pierre  
APPLICANT: Varlet, Paule  
APPLICANT: Vigon, Isabelle  
TITLE OF INVENTION: Polypeptide of a Growth Factor Receptor  
TITLE OF INVENTION: Family, Application in the Diagnosis and Treatment of  
TITLE OF INVENTION: Myeloproliferative Disease  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 90 South 7th Street, 3100 No. 5989833west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,402  
FILING DATE:  
CLASSIFICATION: 436  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,259  
FILING DATE: 20-SEP-1994  
CLASSIFICATION: 436  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/078,311  
FILING DATE: 18-JUN-1993  
CLASSIFICATION: 436  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR90/00762  
FILING DATE: 14-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Randa11 A. Hillson  
REGISTRATION NUMBER: 31,838  
REFERENCE/DOCKET NUMBER: 8076.84US03  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: Extracellular domain of v-mp1  
US-08-460-402-25  
Query Match 3.4%; Score 6; DB 2; Length 32;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 164 LRPRAR 169  
DB 2 LRPRAR 7  
RESULT 4  
US-08-078-311-10  
Sequence 10, Application US/08078311  
Patent No. 5925750  
GENERAL INFORMATION:  
APPLICANT: Charon, Martine  
APPLICANT: Gisselbrecht, Silvie  
APPLICANT: Penciolelli, Jean-Francis  
APPLICANT: Souvri, Michele  
APPLICANT: Tambourin, Pierre  
APPLICANT: Varlet, Paule  
APPLICANT: Vigon, Isabelle  
TITLE OF INVENTION: Polypeptide of a Growth Factor Receptor  
TITLE OF INVENTION: Family, Application in the Diagnosis and Treatment of  
TITLE OF INVENTION: Myeloproliferative Disease  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5925750west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible